现代数学基础丛书 109

矩阵的半张量积
理论与应用

程代展 齐洪胜 著

(O-2722.0101)

销售分类建议：高等数学

定 价：56.00 元

内 容 简 介

本书介绍了一种新的矩阵乘法，称为矩阵的半张量积. 它将矩阵的普通乘法推广到一般情况，即前矩阵的列数与后矩阵的行数不相等的情况. 推广后的乘法仍保持原矩阵乘法几乎所有的性质. 矩阵的半张量积使矩阵方法可以用于处理高维数组及非线性问题. 本书前5章介绍半张量积定义及基本性质，后7章为其各种应用，包括数理逻辑及基于逻辑的智能系统，对微分几何及抽象代数中的一些基本问题的应用，非线性控制系统的镇定，动态系统的对称性，非线性系统的稳定域估计，系统控制中的 Morgan 问题及线性化问题.

本书适合系统科学、控制理论、计算机、人工智能等专业的师生及科研人员阅读参考.

图书在版编目(CIP)数据

矩阵的半张量积:理论与应用/程代展,齐洪胜 著.—北京:科学出版社,2007 现代数学基础丛书,109

ISBN 978-7-03-018833-5

Ⅰ. 矩… Ⅱ. ①程…②齐… Ⅲ. 矩阵-乘法 Ⅳ. 0151.21

中国版本图书馆 CIP 数据核字(2007)第049892号

责任编辑:张 扬/责任校对:赵燕珍

责任印制:赵德静/封面设计:陈 敬

科学出版社 出版

北京东黄城根北街16号

邮政编码: 100717

http://www.sciencep.com

天时彩色印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2007年5月第 一 版 开本:B5(720×1000)

2007年5月第一次印刷 印张:21

印数:1—3 000 字数:392 000

定价:56.00元

(如有印装质量问题，我社负责调换<环伟>)

《现代数学基础丛书》序

对于数学研究与培养青年数学人才而言，书籍与期刊起着特殊重要的作用. 许多成就卓越的数学家在青年时代都曾钻研或参考过一些优秀书籍，从中汲取营养，获得教益.

20 世纪 70 年代后期，我国的数学研究与数学书刊的出版由于“文化大革命”的浩劫已经破坏与中断了十余年，而在这期间国际上数学研究却在迅猛地发展着. 1978 年以后，我国青年学子重新获得了学习、钻研与深造的机会. 当时他们的参考书籍大多还是 50 年代甚至更早期的著述. 据此，科学出版社陆续推出了多套数学丛书，其中《纯粹数学与应用数学专著》丛书与《现代数学基础丛书》更为突出，前者出版约 40 卷，后者则逾 80 卷. 它们质量甚高，影响颇大，对我国数学研究、交流与人才培养发挥了显著效用.

《现代数学基础丛书》的宗旨是面向大学数学专业的高年级学生、研究生以及青年学者，针对一些重要的数学领域与研究方向，作较系统的介绍. 既注意该领域的基础知识，又反映其新发展，力求深入浅出，简明扼要，注重创新.

近年来，数学在各门科学、高新技术、经济、管理等方面取得了更加广泛与深入的应用，还形成了一些交叉学科. 我们希望这套丛书的内容由基础数学拓展到应用数学、计算数学以及数学交叉学科的各个领域.

这套丛书得到了许多数学家长期的大力支持，编辑人员也为其付出了艰辛的劳动. 它获得了广大读者的喜爱. 我们诚挚地希望大家更加关心与支持它的发展，使它越办越好，为我国数学研究与教育水平的进一步提高作出贡献.

杨　乐

2003 年 8 月

前 言

尽管微积分和线性代数在数学中看来是多么初等, 它们却是近代数学的基础, 甚至可以说, 它们是整个自然科学研究中最基本的数字计算的工具. 谁都知道, 牛顿和莱布尼兹发明了微积分. 但是没有人知道谁发明了线性代数. 这也许是因为线性代数和矩阵的概念与线性方程组密切相关, 而线性方程组则是从远古时代人们开始研究的最古老的课题之一. 例如, 在公元前100年出版的《九章算术》一书中, 线性系统的系数就被排成一个方阵的形式, 而西方称之为高斯消元法的方法就被用来求解方程组了. 这里使用的方阵实际上就是矩阵. 一言以蔽之, 矩阵是那样初等和自然, 所以在历史长河中, 许许多多数学家都研究过它, 以至于谁也说不清第一个发明矩阵或线性代数的人是谁.

那么, 矩阵在实际应用中究竟起着什么样的作用呢? 我们引用一本经典的矩阵理论及应用的书中的一段话[40]: "矩阵理论可称为高等算术, 几乎每一个工程应用都涉及到矩阵, 这是因为要处理带有许多部件的复杂系统, 必须有一种数学工具, 它能将这些部件结合在一起. 而矩阵的方法正好能达到这一目的. 在电网络、结构理论、力学系统、经济学研究等中均可找到矩阵理论的精彩应用." 这段话给了矩阵理论一个粗略的描述. 当然, 矩阵也不是万能的, 因此, 我们有必要对目前使用的矩阵工具的能力和适用范围作进一步的分析. 从线性代数知道, 矩阵是处理1维或2维数组的有力工具, 特别是在考虑线性映射或线性函数时, 矩阵是处理这些问题的完美手段. 当我们考虑2维数组时, 用矩阵表示的双线性型或二次型是最有力的工具.

高维数组在数学和工程问题中经常出现, 例如, 在微分几何中我们用称为Christoffel 符号$\{r^i_{jk}\}$ 的3维数组来描述联络[10]. 又如, 要描述n 维空间上的一个协变阶为s、逆变阶为t 的张量, 结构常数$\{w^{i_1,\cdots,i_s}_{j_1,\cdots,j_t}\}$ 是一个$s+t$ 维数组[10]. 再如, 在统计学中, 考虑多变量的协方差时, 要用到高维数组[5, 6]. 一个工程例子是: 在一个化工反应或者一种实验设计中, 如果有k 个因素与产品质量有关, 在做试验时, 每个因素可以取若干不同水平, 那么实验数据就是一个k 维数组.

在考虑高维数组时, 矩阵形式并不方便. 在考虑3维数组时有些人建议用立方阵[5, 6]. 但它并未得到学术界的广泛认可. 其原因首先是它的运算过于复杂, 为了对付不同的情况, 就要定义不同的立方积(参见第 2 章). 其次, 它无法用于更高维(维数>3)的情况, 显然, 将更高维数组排成抽象的更高维的长方体不仅在实现上难以做到, 而且定义运算更加困难.

那么, 怎样才能有效地解决这个问题呢? 计算机科学可以给我们一些启发. 在计算机存储器中数据是排成一个长列, 不管它的维数是多少, 在编程的时候, 例

如在C 语言中, “指针”、“指针的指针”、“指针的指针的指针” 等被用来处理数据的层次. 或者形象地说, 自动寻找高维数组相应的指针所在, 这样, 高维数据就可以被合理地安排和处理了.

受这个启发, 我们提出了一种矩阵乘法, 叫做矩阵的半张量积, 这种乘法的优点就是可以自动地寻找数据的层次, 从而有效地处理多维数组的问题.

从另一个角度看, 向量和二次型都很容易用矩阵处理. 但是, 一般多线性映射就很难用矩阵表示了. 当考虑非线性问题时, 多线性映射是很重要的, 因为多项式就可以由多线性映射组成(严格地说, 齐次多项式是一种特殊的多线性映射, 参见第 5 章). 在半张量积的框架下, 多线性映射很容易用矩阵处理, 而它可以用来逼近一般非线性映射. 所以, 半张量积使我们能用矩阵方法来处理非线性问题. 这是引进半张量积的另一个重要目的.

当然, 结合矩阵普通乘法和Kronecker 积也可以表示多线性映射, 但这两种乘法之间无结合律, 这使得它们在进一步的使用中会有很大困难. 读者通过本书能看到, 在许多场合下, 半张量积可以代替Kronecker 积来表示多线性映射, 而它与普通矩阵乘法之间有结合律, 这为讨论及应用带来了巨大的方便. 实际上, 普通矩阵乘法作为它的一种特例而被自然地吸收到这种新乘法中来了. 正因如此, 它为用矩阵处理多线性映射带来革命性的突破, 使得许多进一步的简化及演变成为可能. 这就是为什么半张量积使非线性问题可以通过矩阵和线性代数的方法得到处理.

以上提到引进这种新矩阵乘法的两个重要动机: 处理多维数组及处理非线性问题. 下面简单介绍一下什么是矩阵半张量积. 大家知道, 数乘和矩阵乘法的第一个巨大不同是前者对所有的数都可乘, 而后者是有条件的: A, B 可乘只有A 的列数(c_A) 与B 的行数(r_B) 相等才行. 为叙述方便, 我们称这个条件为“等维数条件”. 当然, Kronecker 积对任意A, B 都可乘, 但它是另一种乘法, 与普通乘法不同. 简单地说, 半张量积就是把普通矩阵乘法推广到等维数条件不成立, 即$c_A \neq r_B$ 的情况. 虽然我们对任意的A, B 都能定义半张量积, 但目前在应用上真正有意义的是c_A, r_B 中一个为另一个的因子的情况, 称这个条件为“倍维数条件”, 本书主要研究倍维数条件成立的情况.

作为理论上的兴趣, 本书也定义和讨论了任意两个矩阵的半张量乘法, 当两矩阵满足等维数条件时这种乘法与普通乘法一致; 当两矩阵满足倍维数条件时它与上述倍维数条件下定义的乘法一致. 将矩阵乘法推广到任意两个矩阵的情况曾经是第一作者年轻时的梦, 对于这种最一般的情况, 有许多理论研究可以做. 本书之所以未对此展开, 是因为目前尚未发现它有多少实际应用.

矩阵的半张量积首次是作者在文献[17] 中提出, 初步结果被收集在文献[20] 中. 进一步的结果和若干不同的应用发表在文献[21], 文献[27]~[29] 等中.

本书共分 12 章. 第 1 ~ 5 章是关于矩阵半张量积的定义和性质, 其中第 1~3 章讨论左半张量积, 是基础. 本书主要用到的都是左半张量积, 通常我们提到半张量积时实际默认其为左半张量积. 第 4 章考虑右半张量积以及任意两个矩阵的半张量积, 初读时可以略去. 第 5 章讨论多元多项式及其微分的半张量积表示, 它是用半张量积处理非线性问题的基础. 第 6 章到第 12 章为半张量积的各种应用. 第 6 章是它在逻辑及基于逻辑的智能系统中的应用. 第 7 章是它在一些几何、代数及物理问题中的应用, 如微分几何中联络的半张量积表示、代数结构的半张量积公式、相对论中张量场缩并的半张量积证明. 第 8 章是它在控制系统镇定中的应用, 研究如何通过设计中心流形来镇定非最小相位的非线性系统, 半张量积方法将用于多项式反馈设计. 第 9 章考虑动态系统对称性中的应用, 介绍对称性条件的半张量积表达. 第 10 章考虑动态系统吸引域问题, 通过半张量积得到吸引域边界公式的Taylor 展式系数公式. 第 11 章考虑Morgan 问题. 通过半张量积, 问题的可解性转换为代数方程解的存在性. 第 12 章考虑线性化问题, 线性化条件均通过半张量积及其性质求得或表示出.

前5章是矩阵半张量积的基础, 因此, 每章后均附有一定数量的习题, 内容除练习性的问题外, 也包括一些新概念和结果, 它们对理解本书很有帮助, 希望读者也能给予重视.

附录给出一些计算程序, 包括半张量积计算、换位矩阵等. 它们可以在Matlab中直接应用. 这些程序不仅为读者提供了学习, 练习及检验理论结果的方便工具, 也可直接应用于理论研究及实际工程设计.

阅读本书的基本要求是微积分和线性代数. 因此, 理工科大学一年级以上的学生都可以理解本书的主要内容. 但同时, 作为应用的例子, 本书涉及到微分几何、抽象代数、李群、李代数、表示论、非线性控制、统计、对策论等许多领域, 这无疑会给读者带来困难. 笔者除努力将问题的背景作尽可能清楚的说明外, 在每一章的最后有一个注释与参考, 它对涉及的有关学科或背景作一个简明的解释, 以便读者理解或查询有关文献. 当然, 这也不可能完全解决问题, 笔者对读者的最后一个建议是: 不必将每个应用实例都弄明白. 不妨选择与自己专业有关或感兴趣的例子读, 它不会影响对本书主要内容的理解和方法的应用.

笔者相信, 矩阵的半张量积像矩阵本身一样, 是一个极基本又极具普遍意义的有力的工具, 它将在几乎所有的工程和科学问题中找到自己的应用. 同时, 由于它是普通矩阵乘法的推广, 当它被学术界广泛接纳后, 人们就可以彻底忘却普通矩阵乘法了. 由于内容的原创性和笔者的水平, 错误甚至缺陷在所难免, 敬请有关专家和读者们不吝赐教. 更希望有兴趣的读者和我们一起, 进一步完善这一新理论.

“半张量积” 这个全新的概念从最初朦胧的想法到今天大体成形, 中间经历

了10年的努力, 和许多同行朋友的讨论, 甚至是面红耳赤的争辩, 才使它逐步合理起来.这里想提到一些在这方面合作过或给过作者重要批评建议和意见的师长及朋友, 他们有清华大学卢强院士、梅生伟教授、北京大学黄琳院士、瑞典皇家工学院胡晓明教授、香港大学吴复立教授、香港中文大学黄捷教授、香港浸会大学薛伟民教授、香港城市大学冯刚教授、美国华盛顿大学谈自忠教授、B. Ghosh 教授、美国Texas Tech 大学C. Martin 教授、W. Dayawansa 教授、美国Bradley 大学刘江波教授等. 还有一些原博士研究生及博士后, 包括清华大学马进博士、刘峰博士、薛安成博士、系统所张利军博士、董亚莉博士等, 他们均在半张量积及其应用研究中做了许多贡献.

本书的主要内容均先后发表在若干国际、国内杂志上, 大部分内容曾在不同的国际会议上报告过. 部分内容在数学与系统科学研究院相关研究生课程上讲授过. 许多同行, 特别是中国科学院系统所的老师和研究生们提出了许多的批评建议, 使本书内容得以完善. 作者对此深表感谢! 同时, 作者感谢清华大学卢强院士、中国科学院系统所陈翰馥院士, 中国科学院数学与系统科学研究院院长郭雷院士, 感谢他们对本书及本项研究工作的一贯支持.

最后, 本书的出版得到中国科学院科学出版基金的支持, 科学出版社张扬编辑为本书出版做了大量策划和编辑工作, 特在此致谢.

作 者

于中科院系统所

2006 年8 月

符号说明

$M_{m\times n}$	$m\times n$ 矩阵集合
M_n	$n\times n$ 矩阵集合
$V_r(M)$	矩阵的行展开
$V_c(M)$	矩阵的列展开
$\otimes$	矩阵的张量积
$\circ$	矩阵的Hadamard 积
$\ltimes$	矩阵的左半张量积
$\rtimes$	矩阵的右半张量积
$\prec$	升序
$\succ$	降序
$\mathrm{diag}(A_1,\cdots,A_n)$	由块$A_1,\cdots,A_n$ 作为对角元的矩阵
$\mathrm{col}(A_1,\cdots,A_n)$	$(A_1^{\mathrm{T}},\cdots,A_n^{\mathrm{T}})^{\mathrm{T}}$
$\mathrm{tr}(A)$	矩阵A 的迹
$A^{-\mathrm{T}}$	矩阵A 的转置逆
$A^{\mathrm{T}(p\times q)}$	矩阵A 的块转置
$\mathbf{1}_n$	$(\underbrace{1,1,\cdots,1}_{k})^{\mathrm{T}}$
δ_i^n	单位阵I_n 的第i 列
δ_n^i	单位阵I_n 的第i 行
L_A	矩阵A 的Lyapunov 映射
$\sigma(A)$	矩阵A 的特征值集合
$\mathrm{Re}\,\sigma(A)$	矩阵A 的特征值的实部
$[A]$	矩阵A 的等价类
$\wedge$	合取
$\vee$	析取
$\neg$	否定
$\rightarrow$	蕴涵
$\leftrightarrow$	等值
$\Rightarrow$	逻辑推断
$\Leftrightarrow$	逻辑等价
$GL(n,\mathbb{R})$	一般线性群
$gl(n,\mathbb{R})$	一般线性代数

$SO(n,\mathbb{R})$	特殊正交线性群
$so(n,\mathbb{R})$	特殊正交线性代数
$Sp(n,\mathbb{R})$	辛群
$sp(n,\mathbb{R})$	辛代数
$\mathbf{S}_n$	置换群
$[\cdot,\cdot]$	李括号
$\mathrm{ad}_f^k g$	向量场g 对f 的k 次李导数
$L_f^k h$	函数h 对f 的k 次李导数
$T_s^r(V)$	V 上协变阶r 逆变阶s 的张量集合
$\Omega^r(V)$	V 上r 阶反对称协变张量集合
$\Omega(V)$	V 上反对称协变张量集合
$\wedge$	反对称协变张量的外积
$Id(i;k)$	指标集
DF	F 的微分
∇F	F 的梯度
F_*	微分同胚F 的正向导出映射
F^*	微分同胚F 的逆向导出映射
$\nabla_f g$	g 沿f 的联络
$a\%b$	整数相除a/b 的余数
$[a]$	不大于a 的最大整数
$T(N)$	流形N 的切空间
$T^*(N)$	流形N 的余切空间
$C^\infty(N)$	流形N 上的C^∞ 函数集合
$C^\omega(N)$	流形N 上的解析函数集合
$V(N)$	流形N 上的光滑向量场集合
$V^\omega(N)$	流形N 上的解析向量场集合
$\phi_t^X(x_0)$	向量场X 的以x_0 为初值的积分曲线
$\{\cdot,\cdot\}$	Poisson 括号
$\mathrm{Hess}(h(x))$	$h(x)$ 的Hessian 矩阵
$H < G$	H 是G 的子群
$H \lhd G$	H 是G 的正规子群

目　　录

第 1 章　高维数组及其矩阵形式

本章介绍高维数组的排列. 当数组的维数大于1 时, 它可以排列成不同的矩阵形式. 我们首先对多指标集及其刻画的数组的排列形式给出严格的定义, 从而导出指标形式与相应的数组矩阵形式之间的关系. 为了将一种矩阵形式转化为另一种矩阵形式, 引入了两个辅助工具: 矩阵的块转置和换位矩阵. 然后讨论数组的不同排列形式以及它们之间的转化问题. 这些概念和工具在本书以后的讨论中起着重要作用.

1.1　高维数组

在科学计算中一项基本的工作就是处理数据. 我们将一组数据称为数组. 基于数组的自然特性, 通常每个数组都有它的维数. 先让我们来看看什么是数组的维数. 也许不能轻易地给出数组维数的严格定义, 但是粗略地说, 由k 个独立变量产生的数组, 或者说, 从一个k 维空间中采样出的数据就是一个k 维数组. 从应用的观点来看, 大多数情况下数组的维数是显而易见的. 下面我们给出一些简单的例子来描述它.

例 1.1.1　(1) 测量一天的温度. 如果采样周期是5 分钟, 那么得到288 个数据, 可将它们表示成$t_1, t_2, \cdots, t_{288}$, 这是一个1 维数组.

(2) 考虑一个$n\times n$ 矩阵M. 显然它的所有元素$m_{11}, m_{12}, \cdots, m_{nn}$ 是一个2 维数组.

(3) 为了在区域$\{x \in \mathbb{R}^k \,|\, a_i \le x_i \le b_i, i = 1, \cdots, k\}$ 上数值地描述函数$f(x_1, x_2, \cdots, x_k)$, 我们可以将每个区间$[a_i, b_i]$ 分割成$n_i - 1$ 个子区间, 这样得到n_i 个分割点$x_i^1, x_i^2, \cdots, x_i^{n_i}$, 这些分割点共产生$n = \prod_{i=1}^k n_i$ 个结点. 计算函数在这n 个结点的值, 就得到了一个k 维数组

$$\{v_{s_1,s_2,\cdots,s_k} = f(x_1^{s_1}, x_2^{s_2}, \cdots, x_k^{s_k}) \,|\, s_i = 1, 2, \cdots, n_i,\ 1 \le i \le k\}.$$

□

从上面的例子可以看出, 一个数组的维数可以由其独立指标的个数来决定. 本书中若无特别说明仅考虑由指标索引的有限数据集合, 在这种情况下我们给出数组维数的严格定义.

定义 1.1.2　称一个数组是k 维的, 如果它的元素是由k 个指标索引的. 更确

切一些, 称S 是一个k 维数组, 如果它能够表示成

$$S=\{s_{i_1,\cdots,i_k} \mid i_1=1,\cdots,n_1;\ \cdots;\ i_k=1,\cdots,n_k\}. \tag{1.1.1}$$

S 的大小, 即S 中元素个数, 记作$|S|$.

显然, (1.1.1) 中S 的大小为

$$|S|=n_1\times n_2\times\cdots\times n_k.$$

如果数组是1 维的, 可以将它排列成一个向量, 当数组是2 维时可以将它排列成一个矩阵, 但是当数组维数高于2 时, 我们如何排列它呢? 我们不能将它排列成3 维或者更高维数矩阵的形式, 而且即使我们可以将它排列成那种形式, 我们又如何进行计算呢?

对于3 维数组, 一些学者引入了立方阵的概念. 在下一章的讨论中可以看出立方阵的引入并不是很自然, 定义也比较复杂, 而且立方阵乘法与普通矩阵乘法不相容, 这在很大程度上限制了它的应用. 再者, 很难将它推广到更高维数的情形. 本书的目的就是要找到一个统一的且简洁有效的方法来解决高维数组的矩阵表示及其运算的问题.

本书介绍的这种新算法的一个启示来自计算机科学. 计算机科学家为我们提供了一个思路: 在存储器中, 数据不需要排列成高于1 维的形式. 在内存中, 它们排列成一个序列. 但是在程序设计中, 例如在C 语言中, 人们使用所谓的"指针","指针的指针"、"指针的指针的指针"等等来区分数据的不同层次. 因此, 问题不是如何排列这些数据, 而是如何去确定这些高维数组的层次结构. 换句话说, 问题的关键是: 我们要用适当的方法操纵数据而不是排列数据. 这就是本书提出的一种新的矩阵乘法的基本出发点. 在这种新方法中, 索引起着重要的作用. 下面我们给出索引的严格定义.

定义 1.1.3 给定一组含有$\prod_{i=1}^{k} n_i$ 个元素的数组S, 如(1.1.1) 中S 的元素由k 个指标来索引. 将S 排列成一行(或一列), 说它是由指标$i_1,\cdots,i_k$ 按照索引

$$Id(i_1,\cdots,i_k;n_1,\cdots,n_k).$$

排列的, 如果它是由指标$i_1,\cdots,i_k$ 按照如下方式排序: i_t, $t=1,\cdots,k$ 按照字母表次序依次从1 跑到n_t. 更确切地说, $s_{\alpha_1,\cdots,\alpha_k}$ 排在$s_{\beta_1,\cdots,\beta_k}$ 之前, 当且仅当存在$1\le j\le k$, 使得

$$\alpha_i=\beta_i,\ i=1,\cdots,j-1,\quad \alpha_j<\beta_j.$$

如果各指标的取值个数$n_1,\cdots,n_k$ 是显见的, 上述指标索引可简记为

$$Id(i_1,\cdots,i_k):=Id(i_1,\cdots,i_k;n_1,\cdots,n_k).$$

例 1.1.4 (1) 设$S = \{s_{ijk} \mid i=1,2,3; j=1,2; k=1,2\}$. 如果我们按照索引$Id(i,j,k)$ 排列它, 它就变为

$$s_{111}, s_{112}, s_{121}, s_{122}, s_{211}, s_{212}, s_{221}, s_{222}, s_{311}, s_{312}, s_{321}, s_{322}.$$

如果按照索引$Id(j,k,i)$ 排列, 它变为

$$s_{111}, s_{211}, s_{311}, s_{112}, s_{212}, s_{312}, s_{121}, s_{221}, s_{321}, s_{122}, s_{222}, s_{322}.$$

(2) 设$X = \{a_1, a_2, \cdots, a_{24}\}$. 我们用指标$\lambda_1$, λ_2, λ_3 将其表示为$a_i = a_{\lambda_1,\lambda_2,\lambda_3}$ 的形式, 那么,在不同索引下,它们有不同的形式, 例如:

(a) 按照索引$Id(\lambda_1,\lambda_2,\lambda_3; 2,3,4)$, 这些元素排列成

$$\begin{matrix} a_{111} & a_{112} & a_{113} & a_{114} \\ a_{121} & a_{122} & a_{123} & a_{124} \\ a_{131} & a_{132} & a_{133} & a_{134} \\ \vdots & \vdots & \vdots & \vdots \\ a_{231} & a_{232} & a_{233} & a_{234} \end{matrix}$$

(b) 按照索引$Id(\lambda_1,\lambda_2,\lambda_3; 3,2,4)$, 这些元素排列成

$$\begin{matrix} a_{111} & a_{112} & a_{113} & a_{114} \\ a_{121} & a_{122} & a_{123} & a_{124} \\ a_{211} & a_{212} & a_{213} & a_{214} \\ \vdots & \vdots & \vdots & \vdots \\ a_{321} & a_{322} & a_{323} & a_{324} \end{matrix}$$

(c) 按照索引$Id(\lambda_1,\lambda_2,\lambda_3; 3,2,4)$, 这些元素排列成

$$\begin{matrix} a_{111} & a_{112} & a_{113} \\ a_{121} & a_{122} & a_{123} \\ a_{211} & a_{212} & a_{213} \\ \vdots & \vdots & \vdots \\ a_{421} & a_{422} & a_{423} \end{matrix}$$

注 在以上的排列中, 分行是因为空间的限制, 同时也使数组的层次结构表示得比较清楚. 实际上, 应当将各行依次接起来, 变成一行. □

不同的索引, 实际上相当于不同指标的置换. 指标的置换, 则引起数组元素排列即数据的置换. 因此, 置换群在高维数组的排列及其相关运算中起着十分重要的作用. 为方便计, 我们对有限个元素的置换群及其运算作一简单回顾.

设$\mathbf{S}_k$ 是k 阶置换群, 也就是它由k 个元素的所有置换组成. 简单地用$1,\cdots,k$ 表示这些元素, 假如$k=5$, 那么$\mathbf{S}_5$ 表示五个元素$(1,2,3,4,5)$ 的所有可能的置换. 设其中的某个元素$\sigma\in\mathbf{S}_5$ 可表示为

$$\sigma=\begin{bmatrix}1&2&3&4&5\\ \downarrow&\downarrow&\downarrow&\downarrow&\downarrow\\ 2&3&1&5&4\end{bmatrix}\in\mathbf{S}_5,$$

即σ 将1 变成2, 2 变成3, 3 变成1, 4 变成5, 5 变成4. 可以将σ 简记为轮换式形式

$$\sigma=(1,2,3)(4,5).$$

设$\mu\in\mathbf{S}_5$且

$$\mu=\begin{bmatrix}1&2&3&4&5\\ \downarrow&\downarrow&\downarrow&\downarrow&\downarrow\\ 4&3&2&1&5\end{bmatrix},$$

那么$\mathbf{S}_5$ 中两个元素的乘积可以定义为

$$\mu\sigma=\begin{bmatrix}1&2&3&4&5\\ \downarrow&\downarrow&\downarrow&\downarrow&\downarrow\\ 2&3&1&5&4\\ \downarrow&\downarrow&\downarrow&\downarrow&\downarrow\\ 3&2&4&5&1\end{bmatrix},$$

即$\mu\sigma=(1,3,4,5)$.

如果(1.1.1) 中的数组S 是按照索引$Id(i_1,\cdots,i_k)$ 排列的, 就说它是按照自然序排列的. 当然, 它也可以按照索引$(i_{\sigma(1)},\cdots,i_{\sigma(k)})$ 排列, 即首先指标$i_{\sigma(k)}$ 从1 跑到$n_{\sigma(k)}$, 接着$i_{\sigma(k-1)}$ 从1 跑到$n_{\sigma(k-1)}$, 依此下去. 当然按照不同的索引排列, 元素将可能出现在不同的位置上.

定义 1.1.5　设σ 是$\{1,\cdots,k\}$ 的一个置换. 给出一个含有$\prod_{i=1}^k n_i$ 个元素且被排列成一行(或一列) 的数组S, 我们称它是由指标$i_1,\cdots,i_k$ 按照索引$Id(i_{\sigma(1)},\cdots,i_{\sigma(k)};n_{\sigma(1)},\cdots,n_{\sigma(k)})$ 排列的, 如果它由指标$i_1,\cdots,i_k$ 按照下面索引排序: 首先$i_{\sigma(k)}$ 从1 跑到$n_{\sigma(k)}$, 接着$i_{\sigma(k-1)}$ 从1 跑到$n_{\sigma(k-1)}$, 依此下去直到最后$i_{\sigma(1)}$ 从1 跑到$n_{\sigma(1)}$.

下面介绍几个记号, 设$a \in \mathbb{Z}$, $b \in \mathbb{Z}_+$, 与C 语言中一样, 用$a\%b$ 表示a/b 的余数, $[t]$ 表示不大于t 的最大整数. 例如:

$$100\%3 = 1, \quad 100\%7 = 2, \quad (-7)\%3 = 2;$$
$$[7/3] = 2, \quad [-1.25] = -2.$$

容易证明下面这个简单关系式:

$$a = [a/b]\, b + a\%b. \tag{1.1.2}$$

下一步, 我们考虑指标索引转换问题. 有时我们可能需要将一个单指标索引转换成多指标索引, 或者相反. 特别是通过计算机程序构造某些类型的矩阵时, 指标索引的转换会非常有用. 下面给出转换公式, 它们不难通过数学归纳证明, 我们将证明留给读者.

命题 1.1.6 *设S 是一个长度为$n = \Pi_{i=1}^{k} n_i$ 的数组, 如果我们分别将它用单指标索引与k 指标索引$Id(\lambda_1, \cdots, \lambda_k; n_1, \cdots, n_k)$ 排序如下:*

$$S = \{ s_p \mid p = 1, \cdots, n \} = \{ \sigma_{\lambda_1, \cdots, \lambda_k} \mid 1 \le \lambda_i \le n_i;\ i = 1, \cdots, k \},$$

则

(1) 令$p_k := p - 1$, 那么单指标索引p 可以通过下面的递推公式转换成自然序的多指标索引$(\lambda_1, \cdots, \lambda_k)$:

$$\begin{cases} \lambda_k = p_k\%n_k + 1, \\ p_j = [p_{j+1}/n_{j+1}], \quad \lambda_j = p_j\%n_j + 1, \quad j = k-1, \cdots, 1. \end{cases} \tag{1.1.3}$$

(2) 反过来, 从自然序的多指标索引到单指标索引的变换, 我们有:

$$p = \sum_{j=1}^{k-1} (\lambda_j - 1) n_{j+1} n_{j+2} \cdots n_k + \lambda_k. \tag{1.1.4}$$

下面的例子展示了不同类型指标索引之间的转换.

例 1.1.7 回顾例1.1.4 的第二部分, 我们可以通过不同的索引来标注这些元素, 于是有

(1) 按照索引$Id(\lambda_1,\lambda_2,\lambda_3;2,3,4)$ 考虑a_{11}: 利用(1.1.3), 我们有

$$\begin{aligned}
&p_3 = p-1 = 10,\\
&\lambda_3 = p_3\%n_3+1 = 10\%4+1 = 2+1 = 3,\\
&p_2 = [p_3/n_3] = [10/2] = 2,\\
&\lambda_2 = p_2\%n_2+1 = 2\%3+1 = 2+1 = 3,\\
&p_1 = [p_2/n_2] = [2/4] = 0,\\
&\lambda_1 = p_1\%n_1+1 = 0\%2+1 = 1.
\end{aligned}$$

因此$a_{11} = a_{133}$.

(2) 按照索引$Id(\lambda_1,\lambda_2,\lambda_3;2,3,4)$ 考虑a_{214}: 利用(1.1.4), 我们有

$$p = (\lambda_1-1)n_2n_3+(\lambda_2-1)n_3+\lambda_3 = 1\cdot 3\cdot 4+0+4 = 16.$$

因此$a_{214} = a_{16}$.

(3) 按照索引$Id(\lambda_2,\lambda_3,\lambda_1;3,4,2)$, 数据排列成

$$\begin{matrix}
a_{111} & a_{211} & a_{112} & a_{212}\\
a_{113} & a_{213} & a_{114} & a_{214}\\
\vdots & \vdots & \vdots & \vdots\\
a_{131} & a_{231} & a_{132} & a_{232}\\
a_{133} & a_{233} & a_{134} & a_{234}.
\end{matrix}$$

对于这个索引, 如果想在单指标索引和多指标索引之间进行转换, 我们可以按照自然序构造一个正规索引$b_{\Lambda_1,\Lambda_2,\Lambda_3}$, 其中$\Lambda_1 = \lambda_2$, $\Lambda_2 = \lambda_3$, $\Lambda_3 = \lambda_1$; $N_1 = n_2 = 3$, $N_2 = n_3 = 4$, $N_3 = n_1 = 2$. 那么b_{ijk} 由$(\Lambda_1,\Lambda_2,\Lambda_3)$ 按照索引$Id(\Lambda_1,\Lambda_2,\Lambda_3;N_1,N_2,N_3)$ 排列. 现在我们就可以利用(1.1.3) 和(1.1.4) 来转换指标索引形式了.

例如, 考虑a_{124}. 记$a_{124} = b_{241}$, 对b_{241} 应用(1.1.3) 有

$$\begin{aligned}
p &= (\Lambda_1-1)N_2N_3+(\Lambda_2-1)N_3+\Lambda_3\\
&= (2-1)\times 4\times 2+(4-1)\times 2+1 = 8+6+2 = 15.
\end{aligned}$$

因此

$$a_{124} = b_{241} = b_{15} = a_{15}.$$

再考虑a_{17}. 既然$a_{17} = b_{17}$, 利用(1.1.2) 有

$$\begin{aligned}
&p_3 = p-1 = 16, \quad \Lambda_3 = p_3\%N_3+1 = 1,\\
&p_2 = [p_3/N_3] = 8, \quad \Lambda_2 = p_2\%N_2+1 = 1,\\
&p_1 = [p_2/N_2] = 2, \quad \Lambda_1 = p_1\%N_1+1 = 3.
\end{aligned}$$

这样$a_{17} = b_{17} = b_{311} = a_{131}$.　□

我们已经介绍了一个高维数组可以排列成一个向量, 那么矩阵作为一个2 维数组当然也可以排列成一个向量. 考虑矩阵

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

那么A 的行展开, 记为$V_r(A)$, 是通过将矩阵元素一行接着一行排列而成的nm 维向量, 即

$$V_r(A) = (a_{11}, a_{12}, \cdots, a_{1n}, \cdots, a_{m1}, a_{m2}, \cdots, a_{mn})^{\mathrm{T}}. \tag{1.1.5}$$

A 的列展开, 记为$V_c(A)$, 是通过将矩阵元素一列接着一列排列而成的nm 维向量, 即

$$V_c(A) = (a_{11}, a_{21}, \cdots, a_{m1}, \cdots, a_{1n}, a_{2n}, \cdots, a_{mn})^{\mathrm{T}}. \tag{1.1.6}$$

由定义直接得到

命题 1.1.8

$$V_c(A) = V_r(A^{\mathrm{T}}); \quad V_r(A) = V_c(A^{\mathrm{T}}). \tag{1.1.7}$$

1.2　高维数组的矩阵表示

在上一节里介绍了如何将一个高维数组按不同的指标索引排列成向量形式, 但有时将数组排列成矩阵形式将会带来更多便利.

对于一个高维数组, 例如(1.1.1) 中的S, 有很多不同的方式将它表示成一个矩阵形式, 下面我们就来详细说明.

定义 1.2.1

(1) 形如 (1.1.1) 中的一个高维数组 S, 称它是按照索引 $\binom{\sigma(1),\cdots,\sigma(t)}{\sigma(t+1),\cdots,\sigma(k)}$ 排列的, 如果它被排列成一个$n_{\sigma(1)}\cdots n_{\sigma(t)} \times n_{\sigma(t+1)}\cdots n_{\sigma(k)}$ 矩阵, 而且这个矩阵的行是由$i_{\sigma(1)},\cdots,i_{\sigma(t)}$ 按照索引$Id(i_{\sigma(1)},\cdots,i_{\sigma(t)})$ 排序的, 列是由$i_{\sigma(t+1)},\cdots,i_{\sigma(k)}$ 按照索引$Id(i_{\sigma(t+1)},\cdots,i_{\sigma(k)})$ 排序的.

(2) 称S 是按照t-行形式排列的, 如果它是按照$\binom{t}{1,\cdots,t-1,t+1,\cdots,k}$ 形式排列的.

(3) 称S 是按照t-列形式排列的. 如果它是按照$\big(^{1,\cdots,t-1,t+1,\cdots,k}\big)$ 形式排列的

例 1.2.2 给定一个数组

$$S = \{a_{ijkt} \mid i = 1,2; j = 1,2; k = 1,2,3; t = 1,2,3\}.$$

它的$\binom{2,4}{3,1}$ 形式, 记为S_{31}^{24}, 是

$$S_{31}^{24} = \begin{bmatrix} a_{1111} & a_{2111} & a_{1121} & a_{2121} & a_{1131} & a_{2131} \\ a_{1112} & a_{2112} & a_{1122} & a_{2122} & a_{1132} & a_{2132} \\ a_{1113} & a_{2113} & a_{1123} & a_{2123} & a_{1133} & a_{2133} \\ a_{1211} & a_{2211} & a_{1221} & a_{2221} & a_{1231} & a_{2231} \\ a_{1212} & a_{2212} & a_{1222} & a_{2222} & a_{1232} & a_{2232} \\ a_{1213} & a_{2213} & a_{1223} & a_{2223} & a_{1233} & a_{2233} \end{bmatrix},$$

它的3-行形式, 记为S^3, 是

$$S^3 = \begin{bmatrix} a_{1111} & a_{1112} & a_{1113} & a_{1211} & \cdots & a_{2211} & a_{2212} & a_{2213} \\ a_{1121} & a_{1122} & a_{1123} & a_{1221} & \cdots & a_{2221} & a_{2222} & a_{2223} \\ a_{1131} & a_{1132} & a_{1133} & a_{1231} & \cdots & a_{2231} & a_{2232} & a_{2233} \end{bmatrix},$$

它的3-列形式, 记为S_3, 是

$$S_3 = \begin{bmatrix} a_{1111} & a_{1121} & a_{1131} \\ a_{1112} & a_{1122} & a_{1132} \\ \vdots & \vdots & \vdots \\ a_{2213} & a_{2223} & a_{2233} \end{bmatrix}.$$

□

稍后, 我们将会看到t-行形式和t-列形式的重要性. 实际上, t-行形式是让行按照单指标索引, 而t-列形式是让列按照单指标索引. 根据定义立刻得到

命题 1.2.3 *对于一组高维数组S, 它的t-行形式是它t-列形式的转置, 即*

$$S_t = (S^t)^{\mathrm{T}}, \quad S^t = (S_t)^{\mathrm{T}}. \tag{1.2.1}$$

从以上的讨论可以看出,高维数组可以排列成许多不同的矩阵形式, 而指标索引起着关键的作用. 后面将会看到, 我们可以根据需要将高维数组排成不同的矩阵形式以简化运算.

1.3 一些例子

在科学和工程问题中将会经常遇到高维数组, 本节中我们将给出一些有代表性的例子, 这些例子在本书以后的讨论中还会遇到.

下面一个例子是关于李代数的. 李代数是近世代数中的一个基本概念, 它在非线性系统控制理论中起着重要作用.

例 1.3.1 设$\mathcal{G}$ 是一个n 维李代数, 它的一组基底是$\{E_1,E_2,\cdots,E_n\}$. 假设它关于这组基底的结构常数是$\{\, c_{ij}^k \mid i,j,k=1,2,\cdots,n\}$, 即李代数上两个基底向量的李括号运算满足

$$[E_i,E_j]=\sum_{k=1}^{n} c_{ij}^k E_k,\quad i,j=1,2,\cdots,n. \tag{1.3.1}$$

现在我们按照$\binom{13}{2}$ 形式(即2-列形式)将结构常数排列成一个$n^2\times n$ 矩阵, 那么就得到下面的矩阵:

$$M_2^{13}=\begin{array}{c} (1)\quad (2)\quad \cdots\quad (n) \\ \begin{bmatrix} c_{11}^1 & c_{12}^1 & \cdots & c_{1n}^1 \\ c_{11}^2 & c_{12}^2 & \cdots & c_{1n}^2 \\ \vdots & \vdots & & \vdots \\ c_{11}^n & c_{12}^n & \cdots & c_{1n}^n \\ \vdots & \vdots & & \vdots \\ c_{n1}^1 & c_{n2}^1 & \cdots & c_{nn}^1 \\ c_{n1}^2 & c_{n2}^2 & \cdots & c_{nn}^2 \\ \vdots & \vdots & & \vdots \\ c_{n1}^n & c_{n2}^n & \cdots & c_{nn}^n \end{bmatrix}\end{array} \begin{array}{c} (j)/(i,k) \\ (11) \\ (12) \\ \vdots \\ (1n) \\ \vdots \\ (n1) \\ (n2) \\ \vdots \\ (nn) \end{array}. \tag{1.3.2}$$

如果我们将它按照$\binom{3}{12}$ 形式(即3-行形式) 排列, 它就变成

$$M_{12}^3=\begin{bmatrix} c_{11}^1 & c_{12}^1 & \cdots & c_{1n}^1 & \cdots & c_{n1}^1 & c_{n2}^1 & \cdots & c_{nn}^1 \\ c_{11}^2 & c_{12}^2 & \cdots & c_{1n}^2 & \cdots & c_{n1}^2 & c_{n2}^2 & \cdots & c_{nn}^2 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ c_{11}^n & c_{12}^n & \cdots & c_{1n}^n & \cdots & c_{n1}^n & c_{n2}^n & \cdots & c_{nn}^n \end{bmatrix}. \tag{1.3.3}$$

当然我们也可以将它按照别的矩阵形式来排列. 这些矩阵我们统称之为李代数$\mathcal{G}$ 的结构矩阵.

利用结构矩阵, 就可以用矩阵乘法来计算两个向量的李括号了. 例如, $X = \sum_{i=1}^{n} x_i E_i$, $Y = \sum_{i=1}^{n} y_i E_i$, 那么$Z = [X,Y] = \sum_{i=1}^{n} z_i E_i$. 将它们的系数作为向量, 我们简记$X = (x_1, \cdots, x_n)^{\mathrm{T}}$ 等等, 在向量形式下利用(1.3.2) 有

$$Z = (X \otimes I_n) M_2^{13} Y. \tag{1.3.4}$$

另外, 由(1.3.3) 可以得到

$$Z = M_{12}^3 (X \otimes Y). \tag{1.3.5}$$

我们将验证留给读者. □

下面一个例子是关于张量的. 实际上, 张量及流形上的张量场的计算是我们引入矩阵半张量积的一个基本出发点. 张量是一个典型的多线性映射, 直接利用矩阵的张量积(Kronecker 积) 也可以计算张量(或张量场)的值, 但因为张量积与普通积之间不满足结合律, 在许多情况下, 特别是理论研究中会有许多不便. 而当张量用半张量积进行计算时, 这一缺点就被克服了.

例 1.3.2 (1) 设V 是一个n 维向量空间, $\{d_1, \cdots, d_n\}$ 是V 的一组基底, 并且

$$\omega : \underbrace{V \times \cdots \times V}_{r} \to \mathbb{R}$$

是一个多线性映射, 也可以称它是一个r 阶协变张量. 令

$$\omega^{i_1 \cdots i_r} = \omega(d_{i_1}, \cdots, d_{i_r}), \quad i_1, \cdots, i_r = 1, \cdots, n.$$

那么它就是一个r 维数组. 将这个数组按照索引

$$Id(i_1, i_2, \cdots, i_r; \underbrace{n, \cdots, n}_{r})$$

排列成一个行向量, 我们称它是张量ω 的结构矩阵, 记为M_ω, 那么

$$M_\omega = (\omega^{11\cdots 1}, \omega^{11\cdots 2}, \cdots, \omega^{11\cdots r}, \cdots, \omega^{rr\cdots 1}, \omega^{rr\cdots 2}, \cdots, \omega^{rr\cdots r}).$$

用列向量$X = (x_1, \cdots, x_n)^{\mathrm{T}}$ 表示$X = \sum_{t=1}^{n} x_t d_t$, 那么对于$X_1, \cdots, X_r \in V$, 我们有

$$\omega(X_1, \cdots, X_r) = M_\omega (X_1 \otimes \cdots \otimes X_r). \tag{1.3.6}$$

(2) 设V^* 是V 的对偶向量空间, 对于V中给定的一组基底$\{d_1,\cdots,d_n\}$, V^* 中的对偶基底是$\{e_1,\cdots,e_n\}$, 并且

$$\omega:\underbrace{V^*\times\cdots\times V^*}_{s}\to\mathbb{R}$$

是一个多线性映射, 也可以称它是一个s 阶逆变张量. 令

$$\omega_{i_1\cdots i_s}=\omega(e_{i_1},\cdots,e_{i_s}),\quad i_1,\cdots,i_s=1,\cdots,n.$$

那么它们就构成了一个s 维数组. 将这个数组按照索引

$$Id(i_1,i_2,\cdots,i_s;\underbrace{n,\cdots,n}_{s})$$

排列成一个列向量, 我们也称它为张量ω 的结构矩阵, 记为M_ω, 那么

$$M_\omega=(\omega_{11\cdots1},\omega_{11\cdots2},\cdots,\omega_{11\cdots s},\cdots,\omega_{ss\cdots1},\omega_{ss\cdots2},\cdots,\omega_{ss\cdots s})^{\mathrm{T}}.$$

用向量$\sigma=(\sigma_1,\cdots,\sigma_n)$ 表示$\sigma=\sum\limits_{t=1}^{n}\sigma_te_t$, 那么对于$\sigma_1,\cdots,\sigma_s\in V^*$, 我们有

$$\omega(\sigma_1,\cdots,\sigma_s)=(\sigma_1\otimes\cdots\otimes\sigma_s)M_\omega. \tag{1.3.7}$$

(3) 记号同上, 并设$\omega:\underbrace{V\times\cdots\times V}_{r}\times\underbrace{V^*\times\cdots\times V^*}_{s}\to\mathbb{R}$ 是一个多线性映射, 那么称ω 是一个$(r,\ s)$ 型张量, 其中r 是协变阶数, s 是逆变阶数. 将$(r,\ s)$ 型张量的全体记为$T_s^r(V)$, 后面我们将详细讨论它.

记

$$\omega_{j_1\cdots j_s}^{i_1\cdots i_r}=\omega(d_{i_1},\cdots,d_{i_r};e_{j_1},\cdots,e_{j_s}),\quad i_1,\cdots,j_s=1,\cdots,n,$$

那么我们就得到一个$(r+s)$ 维数组. 现在我们想将这个数组的n^{r+s} 个元素按照如下方式排列成一个矩阵: 行是由指标$\{i_1,\cdots,i_r\}$ 标注, 按照索引$Id(i_1,\cdots,i_r;\underbrace{n,\cdots,n}_{r})$ 排列, 列是由指标$\{j_1,\cdots,j_s\}$ 标注, 按照索引$Id(j_1,\ j_2,\cdots,j_s;\underbrace{n,\cdots,n}_{s})$ 排列. 于是得到矩阵M_ω 如下

$$M_\omega=\begin{bmatrix}\omega_{1\cdots11}^{1\cdots11} & \omega_{1\cdots11}^{1\cdots12} & \cdots & \omega_{1\cdots11}^{n\cdots nn}\\ \omega_{1\cdots12}^{1\cdots11} & \omega_{1\cdots12}^{1\cdots12} & \cdots & \omega_{1\cdots12}^{n\cdots nn}\\ \vdots & \vdots & & \vdots\\ \omega_{n\cdots nn}^{1\cdots11} & \omega_{n\cdots nn}^{1\cdots12} & \cdots & \omega_{n\cdots nn}^{n\cdots nn}\end{bmatrix}. \tag{1.3.8}$$

矩阵M_ω 称为张量ω 的结构矩阵, 并且不难验证:

$$\begin{aligned}&\omega(\sigma_1,\cdots,\sigma_s;X_1,\cdots,X_r)\\&=(\sigma_1\otimes\cdots\otimes\sigma_s)M_\omega(X_1\otimes\cdots\otimes X_r).\end{aligned}\tag{1.3.9}$$

我们后面将会用到这个公式. □

下面几个例子来自博弈论, 是关于Nash 均衡的. 这一概念在决策与优化中起重要作用. 我们讨论由策略表形成的矩阵, 为寻找Nash 均衡点, 需要对策略矩阵作各种变换, 而这些变换导致了一些新的矩阵运算或变换形式.

例 1.3.3 在博弈论中, 有k 个参与人的博弈可以描述为

$$G=\{S_1,\cdots,S_k;u_1,\cdots,u_k\},$$

其中S_t 是参与人t 的策略集, 也就是说, 他或她有n_t 个策略, 即

$$S_t=\{s_1^t,\cdots,s_{n_t}^t\},\quad t=1,\cdots,k.$$

$$u_t=u_t(s_1,\cdots,s_k),\quad t=1,\cdots,k;\quad s_i\in S_i,\quad i=1,\cdots,t$$

是参与人t 的支付函数, 即当第i 个参与人采用策略s_i, $i=1,\cdots,k$ 时第t 个参与人的付出.

那么支付$u_t(s_{i_1}^1,\cdots,s_{i_k}^k)$, $1\le i_t\le n_t$, $t=1,\cdots,k$ 形成了一个由指标$i_1,\cdots,i_k,t$ 标注的$k+1$ 维数组, 为统一计, 不妨记

$$i_{k+1}:=t,\quad n_{k+1}:=k.$$

可以将这个数组排列成一个称为策略表的矩阵. □

下面给出两个具体生成策略表的例子.

例 1.3.4 (囚徒的困境) 两个共犯同时被捕后, 如果两个都不招供, 他们将会只被判刑一年; 如果一个招供, 另一个不招供, 则一个可获赦免,另一个将被判刑十年;如果两个都招供, 将都被判刑五年. 现在策略是$G_1=G_2=\{C,NC\}$,(C: 招供,NC:不招供). 那么将得到囚徒1 (P_1) 及囚徒2 (P_2) 的策略表(表1.3.1) 所示. □

表 1.3.1　囚徒的困境

(a) P_1 策略表

$P_1\backslash P_2$	C	NC
C	5　5	10　0
NC	0　10	1　1

(b) P_2 策略表

$P_2\backslash P_1$	C	NC
C	5　5	0　10
NC	10　0	1　1

例 1.3.5 假设有三个参与人的博弈如下: $S_1 = \{s_1^1, s_2^1, s_3^1\}$, $S_2 = \{s_1^2, s_2^2\}$, $S_3 = \{s_1^3, s_2^3\}$. 支付由$Id(i,j,k,t)$ 索引, 那么P_1 的策略表应当是按照$\binom{1}{234}$ 形式(或1-行形式) 排列的, 如表1.3.2 所示.

表 1.3.2 P_1的策略表

$P_1\backslash P_2-P_3$	$s_1^2s_1^3$	$s_1^2s_2^3$	$s_2^2s_1^3$	$s_2^2s_2^3$
s_1^1	3 1 2	−1 2 1	2 −1 3	2 0 −1
s_2^1	2 2 0	1 −1 2	1 1 2	0 1 2
s_3^1	1 3 −1	−2 1 −1	0 2 −1	2 4 3

那么一旦知道其他参与人所采用的策略, 通过表1.3.2就可以很方便地找到最佳策略. 也就是说, 如果第2个参与人采用策略s_1^2, 第3个参与人采用策略s_1^3, 那么从表的第一列可以看出策略s_1^1 是参与人1 的最佳策略.

略去注释, 将P_1 的策略表表示成矩阵形式(记作T_1), 则得

$$T_1 = \begin{bmatrix} 3 & 1 & 2 & -1 & 2 & 1 & 2 & -1 & 3 & 2 & 0 & -1 \\ 2 & 2 & 0 & 1 & -1 & 2 & 1 & 1 & 2 & 0 & 1 & 2 \\ 1 & 3 & -1 & -2 & 1 & -1 & 0 & 2 & -1 & 2 & 4 & 3 \end{bmatrix}. \tag{1.3.10}$$

□

1.4 块转置

从前面的讨论可以看出, 对于一个高维数组, 它可以有不同的矩阵表示. 实际上, 对于一个1 维数组, 我们可以有两种表示: 行形式和列形式; 对于一个2 维数组, 我们可以将它表示成A,A^{T}、$V_r(A)$ 或$V_c(A)$等.

不过不同的矩阵表示有着不同的用处. 例如, 在例1.3.1 中$\binom{3}{12}$ 形式(3-行形式) 就便于我们计算. 不同的矩阵表示对于策略表的计算也很有用. 下面我们继续例1.3.5.

例 1.4.1 一个有k 个参与人的博弈描述为

$$G = \{S_1, \cdots, S_k; u_1, \cdots, u_k\}.$$

称一个策略集$(s_1^*, \cdots, s_k^*)$ 为Nash 均衡, 如果

$$u_t(s_1^*, \cdots, s_{t-1}^*, s_t^*, s_{t+1}^*, \cdots, s_k^*) \geq u_t(s_1^*, \cdots, s_{t-1}^*, s_t, s_{t+1}^*, \cdots, s_k^*), \quad t = 1, \cdots, k. \tag{1.4.1}$$

显然Nash 均衡对所有参与人都是一个“最佳”选择, 如果其他参与人都选择了Nash 均衡策略, 那么剩下的这个也得选择Nash 均衡策略.

注意到, 当目标是使支付最小时, (1.4.1) 中的“$\geq$” 将变为“$\leq$ ”, 例1.3.4 即为这种类型.

现在让我们看看如何找到Nash 均衡.

回忆例1.3.5, 根据按1-行形式排列成的策略表, 即对应于第一参与人(P_1)的策略表, 对于给定的s_2 和s_3, 很容易就找到s_1^* 使得

$$u_1(s_1^*, s_2, s_3) \geq u_1(s_1, s_2, s_3).$$

实际上, 我们只要从每一列中找最大的u_1 就行了. 为方便, 我们只要在T_1 上工作即可. 在T_1 中将最大u_1 值加下划线, 则可得到如下P_1 的最佳策略矩阵:

$$T_1 = \begin{bmatrix} \underline{3} & 1 & 2 & -1 & 2 & 1 & \underline{2} & -1 & 3 & \underline{2} & 0 & -1 \\ 2 & 2 & 0 & \underline{1} & -1 & 2 & 1 & 1 & 2 & 0 & 1 & 2 \\ 1 & 3 & -1 & -2 & 1 & -1 & 0 & 2 & -1 & \underline{2} & 4 & 3 \end{bmatrix}. \tag{1.4.2}$$

□

在上面的例子中我们也可以找到最佳的u_2 和u_3, 只要将1-行形式转化为按2-行形式和3-行形式排列的就行了. 也就是说, 将T_1 转换为相应的T_2 及T_3. 为此需要介绍几个新概念.

定义 1.4.2　*设A 是一个$m \times n$ 矩阵, 其中$m = ps$, $n = qt$. 那么A 可以表示成分块形式*

$$A = \begin{bmatrix} A_{11} & \cdots & A_{1q} \\ \vdots & & \vdots \\ A_{p1} & \cdots & A_{pq} \end{bmatrix}, \quad \text{这里} \quad A_{ij} \in M_{s\times t}.$$

我们定义块转置$\mathrm{T}(s \times t)$ *为*

$$A^{\mathrm{T}(s\times t)} = \begin{bmatrix} A_{11} & \cdots & A_{p1} \\ \vdots & & \vdots \\ A_{1q} & \cdots & A_{pq} \end{bmatrix}. \tag{1.4.3}$$

当$t = s = 1$ 时, $\mathrm{T}(s \times t) = \mathrm{T}$ 就是普通的矩阵转置.

根据定义可立即得到以下的基本关系式:

$$\left(A^{\mathrm{T}(s\times t)}\right)^{\mathrm{T}(s\times t)} = A. \tag{1.4.4}$$

例 1.4.3 考虑一个4×4 矩阵

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}.$$

那么

$$A^{\mathrm{T}(2\times 2)}=\begin{bmatrix} a_{11} & a_{12} & a_{31} & a_{32} \\ a_{21} & a_{22} & a_{41} & a_{42} \\ a_{13} & a_{14} & a_{33} & a_{34} \\ a_{23} & a_{24} & a_{43} & a_{44} \end{bmatrix};$$

$$A^{\mathrm{T}(1\times 2)}=\begin{bmatrix} a_{11} & a_{12} & a_{21} & a_{22} & a_{31} & a_{32} & a_{41} & a_{42} \\ a_{13} & a_{14} & a_{23} & a_{24} & a_{33} & a_{34} & a_{43} & a_{44} \end{bmatrix}, \quad \cdots.$$

□

与普通的矩阵转置类似, 对于块转置我们有下面的乘法法则:

命题 1.4.4 设$A\in M_{m\times n}$, α,β 分别是m,n 的因子, 则

$$A^{\mathrm{TT}(\beta\times\alpha)\mathrm{T}}=A^{\mathrm{T}(\alpha\times\beta)}. \tag{1.4.5}$$

证明 利用块转置法则, 通过直接计算即可证明结论. □

例 1.4.5 回忆例 1.3.3. 记 T_s 为关于参与人 s 的策略矩阵, $s=1,\cdots,k$. 设 T_1 已得到. 它应是一个$n_1\times(n_2\cdots n_k k)$ 矩阵. 它的列依索引$Id(i_2,\cdots,i_{k+1};n_2,\cdots,n_k,k)$ 排列. 一般地说, T_s 应是一个$n_s\times(n_1\cdots n_{s-1}n_{s+1}\cdots n_k k)$ 矩阵. 它的列依索引$Id(i_1,\cdots,i_{s-1}i_{s+1}\cdots i_{k+1};n_1,\cdots,n_{s-1},n_{s+1},\cdots,n_k,k)$ 排列.

下面的公式给出由T_1 计算T_2 的方法, 我们将它的证明留给读者:

$$T_2=T_1^{\mathrm{T}(1\times n_3\cdots n_k k)}. \tag{1.4.6}$$

□

例 1.4.6 继续例1.4.1, 为了得到第2 个参与人的策略矩阵, 利用公式可知, 只要将T_1 作转置T(1×6) 即得. 然后, 在T_2 中一列接着一列地选优, 即可得到第2

个参与人的最佳策略矩阵. 同样, 对每个最佳策略, 我们也给它们都加了下划线:

$$T_2 = (T_1)^{\mathrm{T}(1\times 6)}$$
$$= \begin{bmatrix} \underline{3} & \underline{1} & 2 & -1 & \underline{2} & 1 & 2 & \underline{2} & 0 & \underline{1} & -1 & 2 & 1 & \underline{3} & -1 & -2 & 1 & -1 \\ \underline{2} & -1 & 3 & \underline{2} & 0 & -1 & 1 & 1 & 2 & 0 & \underline{1} & 2 & 0 & 2 & -1 & \underline{2} & \underline{4} & 3 \end{bmatrix}. \tag{1.4.7}$$

□

为了得到第3 个参与人的最优策略矩阵, 我们需要“换位矩阵”的概念, 这将在下一节讨论. 因此, 下一节会继续讨论上述例子.

接下来我们考虑将一个矩阵的展开转换回矩阵的问题. 这里我们先介绍两个要用到的记号.

定义 1.4.7 设$x \in \mathbb{R}^N$, 且$N = m \times n$, 我们定义V_r 的逆为

$$V_r^{-1}(x,n) = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ x_{n+1} & x_{n+1} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{(m-1)n+1} & x_{(m-1)n+2} & \cdots & x_{mn} \end{bmatrix}, \tag{1.4.8}$$

定义V_c 的逆为

$$V_c^{-1}(x,m) = \begin{bmatrix} x_1 & x_{m+1} & \cdots & x_{(n-1)m+1} \\ x_2 & x_{m+2} & \cdots & x_{(n-1)m+2} \\ \vdots & \vdots & & \vdots \\ x_m & x_{2m} & \cdots & x_{mn} \end{bmatrix}. \tag{1.4.9}$$

特别, 当$N = n^2$ 时, 可以将x 转换回一个方阵. 我们简记

$$V_r^{-1}(x) := V_r^{-1}(x,n), \quad V_c^{-1} := V_c^{-1}(x,n).$$

为了方便, 如果$N = n \times m + r$, 其中$0 < r < n$ $(0 < r < m)$, 那么我们在$\{x_i\}$后面添加$n - r$ $(m - r)$ 个零形成一个新序列x^e, 即

$$x_i^e = \begin{cases} x_i, & i \le N; \\ 0, & N < i \le N + (n - r), \quad (\text{或} \quad N + (m - r)). \end{cases}$$

定义V_r 的逆为

$$V_r^{-1}(x,n) := V_r^{-1}(x^e,n), \tag{1.4.10}$$

定义V_c 的逆为

$$V_c^{-1}(x,m) := V_c^{-1}(x^e,m). \tag{1.4.11}$$

同理, 当$N=n^2+r$, $0<r\leq 2n$时, 考虑$V^{-1}(x)$, 在它后面添加$2n+1-r$ 个零, 可以将它转化为一个$(n+1)\times(n+1)$ 方阵.

例 1.4.8 设$\{x_i\}=(1\ \ 2\ \ -1\ \ 2\ \ 3\ \ 4)$, 那么

$$V_r^{-1}(x,3)=\begin{bmatrix}1&2&-1\\2&3&4\end{bmatrix},\quad V_c^{-1}(x,3)=\begin{bmatrix}1&2\\2&3\\-1&4\end{bmatrix},$$

$$V_r^{-1}(x,4)=\begin{bmatrix}1&2&-1&2\\3&4&0&0\end{bmatrix},\quad V_r^{-1}(x)=\begin{bmatrix}1&2&-1\\2&3&4\\0&0&0\end{bmatrix}.$$

□

实际上, 块转置也可以实现数组表示形式之间的转换.

命题 1.4.9 设$A\in M_{m\times n}$, 那么

$$(V_c(A))^{\mathrm{T}(m\times 1)}=A,\quad \left(V_r^{\mathrm{T}}(A)\right)^{\mathrm{T}(1\times n)}=A. \tag{1.4.12}$$

最后让我们看看一个数组表示形式之间的转换在统计学中的应用.

例 1.4.10 考虑多维映射的最小二乘回归. 设$\Phi:\mathbb{R}^m\to\mathbb{R}^n$ 是一个多维映射, 假设它是一个线性映射, 并且可以表示成

$$Y=AX+B,\quad X\in\mathbb{R}^m,\quad Y\in\mathbb{R}^n. \tag{1.4.13}$$

$\left\{(x^i,y^i)\,|\,i=1,\cdots,N\right\}$ 是观测数据. 我们通过极小化平方误差

$$E=\sum_{i=1}^{N}(y^i-Ax^i-b)^{\mathrm{T}}(y^i-Ax^i-b) \tag{1.4.14}$$

来估计参数A 和b. 记$\xi=V_r(A)$, 然后令

$$\frac{\partial E}{\partial b}=-2\sum_{i=1}^{N}(y^i-Ax^i-b)=-2\left(\sum_{i=1}^{N}y^i-A\sum_{i=1}^{N}x^i-Nb\right):=0,$$

得到

$$b=\frac{1}{N}\left(\sum_{i=1}^{N}y^i-A\sum_{i=1}^{N}x^i\right). \tag{1.4.15}$$

并通过直接计算有

$$\frac{\partial E}{\partial \xi^{\mathrm{T}}}=-2\sum_{i=1}^{N}(y^i-Ax^i-b)^{\mathrm{T}}\left(I_n\otimes(x^i)^{\mathrm{T}}\right).$$

再令

$$\begin{aligned}\frac{\partial E}{\partial \xi}&=-2\sum_{i=1}^{N}\left(I_n\otimes x^i\right)\left(y^i-Ax^i-b\right)\\&=-2\sum_{i=1}^{N}\left(I_n\otimes x^i\right)y^i+2\sum_{i=1}^{N}\left(I_n\otimes x^i\right)Ax^i\\&\quad+2\sum_{i=1}^{N}\left(I_n\otimes x^i\right)b:=0.\end{aligned} \tag{1.4.16}$$

□

为了简化上面这个表达式, 我们需要下面的引理. 由于证明很容易, 所以把它留给读者.

引理 1.4.11　设$A\in M_{n\times m}$, $X\in\mathbb{R}^m$, 则

$$AX=(I_n\otimes X)V_r(A). \tag{1.4.17}$$

现在对于(1.4.16) 中的第一项, 我们有

$$\left(I_n\otimes x^i\right)y^i=\left(y^i\otimes x^i\right).$$

对于(1.4.16) 中的第二项, 利用引理1.4.11 有

$$\begin{aligned}\left(I_n\otimes x^i\right)Ax^i&=\left(I_n\otimes x^i\right)\left(I_n\otimes(x^i)^{\mathrm{T}}\right)\xi\\&=\left(I_n\otimes[x^i(x^i)^{\mathrm{T}}]\right)\xi.\end{aligned}$$

将(1.4.15) 代入(1.4.16) 的第三项, 再利用引理1.4.11 有

$$\left(I_n \otimes x^i\right) b = \frac{1}{N}\left(I_n \otimes x^i\right)\left(\sum_{j=1}^{N} y^j - A\sum_{j=1}^{N} x^j\right)$$

$$= \frac{1}{N}\left(\sum_{j=1}^{N} y^j \otimes x^i\right) - \frac{1}{N}\left(I_n \otimes \left[\sum_{j=1}^{N} x^i (x^j)^{\mathrm{T}}\right]\right)\xi.$$

于是(1.4.14) 变为$\Psi\xi = \eta$, 其中

$$\begin{cases}\Psi = I_n \otimes \left(\sum\limits_{i=1}^{N}[x^i(x^i)^{\mathrm{T}}] - \frac{1}{N}\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N}[x^i(x^j)^{\mathrm{T}}]\right), \\ \eta = \sum\limits_{i=1}^{N} y^i \otimes x^i - \frac{1}{N}\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N}[y^j \otimes x^i].\end{cases} \tag{1.4.18}$$

假设Ψ 是非异阵, 由

$$\xi = \Psi^{-1}\eta \quad \Rightarrow \quad \hat{A} = V_r^{-1}(\xi) = (\xi^{\mathrm{T}})^{\mathrm{T}(n\times 1)}, \tag{1.4.19}$$

并且

$$\hat{b} = \frac{1}{N}\left(\sum_{i=1}^{N} y^i - \hat{A}\sum_{i=1}^{N} x^i\right). \tag{1.4.20}$$

1.5 换位矩阵

定义 1.5.1 换位矩阵$W_{[m,n]}$ 是一个$mn\times mn$ 矩阵, 定义如下: 它的行和列都是由双指标(i,j) 标注, 列是按照索引$Id(i,j;m,n)$ 排列, 行是按照索引$Id(j,i;n,m)$ 排列, 并且位于$[(I,J),(i,j)]$ 上的元素的值为

$$w_{(IJ),(ij)} = \delta_{i,j}^{I,J} = \begin{cases}1, & I = i \text{ 且 } J = j; \\ 0, & \text{其他}.\end{cases} \tag{1.5.1}$$

稍后将会看到换位矩阵在处理多指标数组时起着非常重要的作用.

例 1.5.2 (1) 设$m = 2$, $n = 3$, 我们按如下方式构造换位矩阵$W_{[m,n]}$:由双指标(i,j) 同时标注行和列, W 的列按照索引$Id(i,j;2,3)$ 排列, 即为(11, 12, 13, 21, 22, 23), W 的行按照索引$Id(j,i;3,2)$ 排列, 即为(11, 21, 12, 22 , 13 , 23). 根

据(1.5.1) 有

$$W_{[2,3]} = \begin{array}{cc} \begin{array}{cccccc} (11)&(12)&(13)&(21)&(22)&(23) \end{array} & \\ \begin{bmatrix} 1&0&0&0&0&0\\ 0&0&0&1&0&0\\ 0&1&0&0&0&0\\ 0&0&0&0&1&0\\ 0&0&1&0&0&0\\ 0&0&0&0&0&1 \end{bmatrix} & \begin{array}{c} (11)\\ (21)\\ (12)\\ (22)\\ (13)\\ (23) \end{array} \end{array}.$$

(2) 考虑$W_{[3,2]}$, 它的列按照索引$Id(i,j;3,2)$ 排列, 行按照索引$Id(j,i;2,3)$ 排列. 于是可得

$$W_{[3,2]} = \begin{array}{cc} \begin{array}{cccccc} (11)&(12)&(21)&(22)&(31)&(32) \end{array} & \\ \begin{bmatrix} 1&0&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&0&1&0\\ 0&1&0&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&0&1 \end{bmatrix} & \begin{array}{c} (11)\\ (21)\\ (31)\\ (12)\\ (22)\\ (32) \end{array} \end{array}.$$

□

换位矩阵的结构具有一些特殊的性质. 通过直接计算可以得到

命题 1.5.3 (1) 换位矩阵的逆和转置为

$$W_{[m,n]}^{\mathrm{T}} = W_{[m,n]}^{-1} = W_{[n,m]}. \tag{1.5.2}$$

(2) 当$m=n$ 时, (1.5.2) 变为

$$W_{[n,n]} = W_{[n,n]}^{\mathrm{T}} = W_{[n,n]}^{-1}. \tag{1.5.3}$$

(3)

$$W_{[1,n]} = W_{[n,1]} = I_n. \tag{1.5.4}$$

$m=n$ 的情形特别重要, 以后将会经常遇到. 为了简化符号, 我们记

$$W_{[n]} := W_{[n,n]}.$$

从(1.5.2) 可以看出$W_{[m,n]}$ 总是一个正交矩阵. 这是因为当将它看作一个线性变换时, 它只是改变向量元素的顺序, 而不改变向量的范数.

换位矩阵还可以实现一个矩阵的行展开和列展开之间的转换:

命题 1.5.4 设$A \in M_{m\times n}$, 那么

$$\begin{cases} W_{[m,n]}V_r(A) = V_c(A), \\ W_{[n,m]}V_c(A) = V_r(A). \end{cases} \tag{1.5.5}$$

证明 注意到$V_r(A)$ 中的元素a_{ij} 是按照索引$Id(i,j;m,n)$ 排列的, 而$V_c(A)$ 中的元素是按照索引$Id(j,i;n,m)$ 排列的. 既然$W_{[m,n]}$ 的列由$Id(i,j;m,n)$ 索引, 行由$Id(j,i;n,m)$ 索引, 通过$W_{[m,n]}$ 的作用, 它将按照索引$Id(i,j;m,n)$ 排列的$V_r(A)$ 中的位于(i,j) 的元素移动到按照索引$Id(j,i;n,m)$ 排列的$V_r(A)$ 中位于(i,j) 的元素, 也就是按照索引$Id(i,j;m,n)$ 排列的$V_r(A^{\mathrm{T}})$ 中位于(j,i) 的元素, 即有

$$W_{[m,n]}V_r(A) = V_r(A^{\mathrm{T}}).$$

再由(1.1.7) 就得到了第一个等式. 在第一个等式两边同乘以$W_{[n,m]}$ 就得到第二个等式. □

对于一个双指标数组$\{a_{ij}\}$, 如果它是按照索引$Id(i,j;m,n)$ 排列的, 那么线性映射$W_{[m,n]}$ 将它变成按照索引$Id(j,i;n,m)$ 排列, 反之也成立. 这就是"换位"的意义所在. 由于换位矩阵非常重要, 后面我们会将它推广到多指标情形. 我们在这里将换位矩阵的这种作用叙述成一个推论的形式.

推论 1.5.5 设一个数组$\{a_{ij}\,|\,1 \le i \le m,\ 1 \le j \le n\}$ 按照索引$Id(i,j;m,n)$ 排列成向量X, 那么

$$Y = W_{[m,n]}X$$

是$\{a_{ij}\}$ 按照索引$Id(j,i;n,m)$ 排列成的向量.

例 1.5.6 (1) 设$X = (x_{11}, x_{12}, x_{13}, x_{21}, x_{22}, x_{23})$, 显然$\{x_{ij}\}$ 是按照索引$Id(i,j;2,3)$ 排列的. 直接计算得

$$Y = W_{[23]}X = \begin{bmatrix} 1&0&0&0&0&0\\ 0&0&0&1&0&0\\ 0&1&0&0&0&0\\ 0&0&0&0&1&0\\ 0&0&1&0&0&0\\ 0&0&0&0&0&1 \end{bmatrix} \begin{bmatrix} x_{11}\\ x_{12}\\ x_{13}\\ x_{21}\\ x_{22}\\ x_{23} \end{bmatrix} = \begin{bmatrix} x_{11}\\ x_{21}\\ x_{12}\\ x_{22}\\ x_{13}\\ x_{23} \end{bmatrix}.$$

即在Y 中元素x_{ij} 是按照索引$Id(j,i;3,2)$ 排列的.

(2) 设$X=(x_1,x_2,\cdots,x_m)^{\mathrm{T}}\in\mathbb{R}^m$, $Y=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}\in\mathbb{R}^n$, 那么有

$$X\otimes Y=(x_1y_1,x_1y_2,\cdots,x_1y_n,\cdots,x_my_1,x_my_2,\cdots,x_my_n)^{\mathrm{T}},$$

$$\begin{aligned}Y\otimes X&=(y_1x_1,y_1x_2,\cdots,y_1x_m,\cdots,y_nx_1,y_nx_2,\cdots,y_nx_m)^{\mathrm{T}}\\&=(x_1y_1,x_2y_1,\cdots,x_my_1,\cdots,x_1y_n,x_2y_n,\cdots,x_my_n)^{\mathrm{T}}.\end{aligned}$$

它们都是由$\{x_iy_j\}$ 组成的. 但是$X\otimes Y$ 中元素是按照索引$Id(i,j;m,n)$ 排列的, 而$Y\otimes X$ 中元素是则按照索引$Id(j,i;n,m)$ 排列的. 根据推论1.5.5, 有

$$Y\otimes X=W_{[m,n]}(X\otimes Y).\tag{1.5.6}$$

□

下面的命题是定义的直接推论.

命题 1.5.7 (1) 设$X=(x_{ij})$ 是一个数组按照索引$Id(i,j;m,n)$ 排列成的列向量, 那么$W_{[m,n]}X$ 是原数组的一个按照索引$Id(j,i;n,m)$ 排列的列向量.

(2) 设$\omega=(\omega_{ij})$ 是一个数组按照索引$Id(i,j;m,n)$ 排列成的行向量, 那么$\omega W_{[n,m]}$ 是原数组的一个按照索引$Id(j,i;n,m)$ 排列的行向量.

换位矩阵同样可以用于多指标数组中某两个指标的排序交换, 它可以看作上述命题的一个推论. 因为它在应用上十分重要, 我们将其叙述成一个定理的形式.

定理 1.5.8 (1) 设 $X=(x_{i_1,\cdots,i_k})$ 是一个数组按照索引 $Id(i_1,\cdots,i_k;n_1,\cdots,n_k)$ 排列成的列向量, 那么

$$\left[I_{n_1+\cdots+n_{t-1}}\otimes W_{[n_t,n_{t+1}]}\otimes I_{n_{t+2}+\cdots+n_k}\right]X$$

是原数组的一个按照索引$Id(i_1,\cdots,i_{t+1},i_t,\cdots,t_k;n_1,\cdots,n_{t+1},n_t,\cdots,n_k)$ 排列的列向量.

(2) 设 $\omega=(\omega_{i_1,\cdots,i_k})$ 是一个数组按照索引

$$Id(i_1,\cdots,i_k;n_1,\cdots,n_k)$$

排列成的行向量, 那么

$$\omega\left[I_{n_1+\cdots+n_{t-1}}\otimes W_{[n_{t+1},n_t]}\otimes I_{n_{t+2}+\cdots+n_k}\right]$$

是原数组的一个按照索引

$$Id(i_1,\cdots,i_{t+1},i_t,\cdots,t_k;n_1,\cdots,n_{t+1},n_t,\cdots,n_k)$$

排列的行向量.

证明 对于列向量, 根据定义, 结论是显见的. 对于行向量, 只要取一转置即得. □

例 1.5.9 继续例1.4.1 和例1.4.6. 为了得到按3-形式排列的(即参与人P_3) 的最优策略矩阵, 我们可以从T_1 出发. 首先将参与人j 和k 交换位置, 注意在T_1 中每一行的元素是按$Id(j,k,t;2,2,3)$ 排列的. 由定理1.5.8, 要让它依$Id(k,j,t;2,2,3)$ 排列, 应右乘$W_{[2]}\otimes I_3$. 现在, 指标k 已换到第一位, 于是, 按照上一节(T_2) 的构造方法, 再作一次块转置$\mathrm{T}(1\times n_2k)=\mathrm{T}(1\times 6)$ 即得. 于是有

$$\mathrm{T}_3=\left[\mathrm{T}_1(W_{[2]}\otimes I_3)\right]^{\mathrm{T}(1\times 6)}=$$
$$\begin{bmatrix} \underline{3} & \underline{1} & 2 & \underline{2} & -1 & 3 & 2 & \underline{2} & 0 & 1 & 1 & 2 & 1 & \underline{3} & -1 & 0 & 2 & -1 \\ -1 & \underline{2} & 1 & \underline{2} & 0 & -1 & \underline{1} & -1 & 2 & 0 & \underline{1} & 2 & -2 & 1 & -1 & \underline{2} & \underline{4} & 3 \end{bmatrix} \tag{1.5.7}$$

于是可在P_3 的最佳策略矩阵中一列接着一列地找到最佳策略s_3, 然后也用下划线标出. 为了让初学者看得更清楚, 我们把它还原成策略表:

最后, 从表1.5.1 中可以看出它有两个Nash 均衡点: (s_1^1,s_1^2,s_1^3) 和(s_3^1,s_2^2,s_2^3).

表 1.5.1 标出最佳u_3的策略表

$P_3\backslash P_1-P_2$	$s_1^1s_1^2$	$s_1^1s_2^2$	$s_2^1s_1^2$	$s_2^1s_2^2$	$s_3^1s_1^2$	$s_3^1s_2^2$
s_1^3	$\underline{3},\underline{1},\underline{2}$	$\underline{2},-1,\underline{3}$	$2,\underline{2},0$	$1,1,\underline{2}$	$1,\underline{3},\underline{-1}$	$0,2,-1$
s_2^3	$-1,\underline{2},1$	$\underline{2},0,-1$	$\underline{1},-1,\underline{2}$	$0,\underline{1},\underline{2}$	$-2,1,\underline{-1}$	$\underline{2},\underline{4},\underline{3}$

□

最后, 让我们看看如何用计算的方法构造$W_{[m,n]}$.

换位矩阵可以按照下面的方法来构造: 记$\mathbb{R}^n$ 的标准基的第i 个为δ_i^n, 实际上δ_i^n 就是单位阵I_n 的第i 列, 那么有

命题 1.5.10

$$W_{[m,n]}=\left[\delta_1^n\ltimes\delta_1^m \quad \cdots \quad \delta_n^n\ltimes\delta_1^m \quad \cdots \quad \delta_1^n\ltimes\delta_m^m \quad \cdots \quad \delta_n^n\ltimes\delta_m^m\right]. \tag{1.5.8}$$

证明 利用(1.5.5), 有

$$W_{[m,n]}\delta_i^m\delta_j^n=\delta_j^n\delta_i^m.$$

也就是说, 如果$W_{[m,n]}$ 的列由双指标(i,j) 按照次序$11,\cdots,1n,\ \cdots,\ m1,\cdots,mn$ 索引的, 那么它的第(i,j) 列就是$\delta_j^n\delta_i^m$, 这样就得到(1.5.8). □

为了便于以后应用, 我们在这里给出换位矩阵的其他两种表示形式, 并将证明作为练习留给读者.

记I_n 的第i 行为δ_n^i, 那么有

$$W_{[m,n]} = \begin{bmatrix} I_m \otimes \delta_n^1 \\ \vdots \\ I_m \otimes \delta_n^n \end{bmatrix}. \tag{1.5.9}$$

类似的我们还有

$$W_{[m,n]} = \left[I_n \otimes \delta_1^m, \quad \cdots, \quad I_n \otimes \delta_m^m \right]. \tag{1.5.10}$$

最后, 我们指出, 换位矩阵具有因式分解的形式.

命题 1.5.11　换位矩阵具有如下的分解形式

$$W_{[p,qr]} = (I_q \otimes W_{[p,r]})(W_{[p,q]} \otimes I_r) = (I_r \otimes W_{[p,q]})(W_{[p,r]} \otimes I_q), \tag{1.5.11}$$

$$W_{[pq,r]} = (W_{[p,r]} \otimes I_q)(I_p \otimes W_{[q,r]}) = (W_{[q,r]} \otimes I_p)(I_q \otimes W_{[p,r]}). \tag{1.5.12}$$

证明　令$X_1 \in \mathbb{R}^p$, $X_2 \in \mathbb{R}^q$ 及$X_3 \in \mathbb{R}^r$, 则

$$W_{[p,qr]} X_1 X_2 X_3 = X_2 X_3 X_1,$$

同时

$$(W_{[p,q]} \otimes I_r) X_1 X_2 X_3 = X_2 X_1 X_3,$$

$$(I_q \otimes W_{[p,r]})(W_{[p,q]} \otimes I_r) X_1 X_2 X_3 = (I_q \otimes W_{[p,r]}) X_2 X_1 X_3 = X_2 X_3 X_1.$$

即

$$W_{[p,qr]} X_1 X_2 X_3 = (I_q \otimes W_{[p,r]})(W_{[p,q]} \otimes I_r) X_1 X_2 X_3.$$

由于X_1, X_2 和X_3 是任选的, 这就证明了(1.5.11) 的第一式. 交换q, r 即得(1.5.11) 的第二式.

(1.5.12) 的证明类似.　□

1.6　注释与参考

本节给出本章涉及到的一些数学概念的简单说明, 并指出有关参考文献.

关于立方阵的概念可参考文献[107], [5], [97], [108]. 我们将在下一章进一步讨论它们, 并说明立方阵的各种乘法都不过是半张量积的特例.

置换群实际上可以看作一个集合到自身的可逆变换的全体，其运算为映射的复合. 本书只考虑有限集合的情况. 本章对此作了简单介绍，更详细的论述可见任一近世代数教程，例如文献[67].

李代数首先是一个向量空间，同时，向量间定义了一个双线性的"乘法"（称为李括号)，它满足反对称性和Jacobi 恒等式. 关于李代数的概念可参见第7 章，更详细的内容可见一些标准教科书，例如文献[10].

张量可以看作对一组向量空间和一组余向量空间的多线性映射. 在一个微分流形上可以定义一个张量场，它在流形的每一点就是其切空间的乘积空间和余切空间乘积空间上的多线性映射. 关于张量及张量场的基本概念和计算可参见一些标准教科书，例如文献[10], [92].

Nash 均衡(Nash equilibrium) 在现代经济学中起着重要作用，Nash 也因此获得了诺贝尔经济学奖，有关的基本概念可参考文献[78].

换位矩阵最早是在文献[63] 中出现的，在那里它是用(1.5.10) 形式定义的，称为交换矩阵(commutation matrix). 虽然和这里的定义有所不同，但是它们在本质上却是一样的. 本书的定义有更明确的物理意义，从这个定义出发,我们证明了它的一系列性质.

习 题 一

1.1 证明(1.1.2) 和(1.1.3).

1.2 验证(1.3.4) 和(1.3.5).

1.3 检验公式(1.4.5)的正确性.

1.4 检验关系式(1.4.6)的正确性.

1.5 设V 是一个 n 维向量空间，它的一组基底是$\{d_1,\cdots,d_n\}$，对偶基底是$\{e_1,\cdots,e_n\}$. 考虑张量$\omega\in T_s^r(V)$,在这组基底下ω的结构矩阵是$M_{\pi\omega}$. 设$\{\tilde{d}_1,\cdots,\tilde{d}_n\}$是$V$ 的另一组基底，并且有

$$\begin{bmatrix}\tilde{d}_1\\ \vdots\\ \tilde{d}_n\end{bmatrix}=A\begin{bmatrix}d_1\\ \vdots\\ d_n\end{bmatrix}.$$

求ω 在新基底下的结构矩阵.

1.6 行向量$X=(x_{i,j})\in\mathbb{R}^{pq}$ 由(i,j) 按照索引$Id(i,j;p,q)$ 排列. 求矩阵Φ 使得$Y=X\Phi$，其中Y 是由相同的元素按照索引$Id(j,i;q,p)$ 排列的.

1.7 行向量$X=(x_{i_1,\cdots,i_k})$ 由$(i_1,\cdots,i_k)$ 按照索引

$$Id(i_1,\cdots,i_p,\cdots,i_q,\cdots,i_k;i_1,\cdots,n_p,\cdots,n_q,\cdots,n_k)\quad(p<q)$$

排列. 求矩阵Φ 使得$Y = X\Phi$, 其中Y 是由相同的元素按照索引

$$Id(i_1,\cdots,i_q,\cdots,i_p,\cdots,i_k;i_1,\cdots,n_q,\cdots,n_p,\cdots,n_k)$$

排列的.

1.8 给出一个s 维数组

$$A=\{a_{i_1\cdots i_s}\,|\,i_k=1,\cdots,n_k,\quad k=1,\cdots,s\}.$$

(1) A_p和A_q 分别表示A 的p-行形式和q-行形式, 其中$1\le p<q\le s$. 请用A_q 表示出A_p.

(2) A^p 和A^q 分别表示A的p-列形式和q-列形式, 其中$1\le p<q\le s$. 请用A^q 表示出A^p.

(提示: 利用习题1.6 的结果和块转置)

1.9 给定数组$S = \{s_{i_1,\cdots,i_k}\,|\,1\le it\le n_t; t=1,\cdots,k\}$. 令$S_1$ 为S 依$\binom{\sigma(1),\cdots,\sigma(t)}{\sigma(t+1),\cdots,\sigma(k)}$排列, 令$S_2$ 为S 依$\binom{\sigma(t+1),\cdots,\sigma(k)}{\sigma(1),\cdots,\sigma(t)}$ 排列. 证明

$$S_1=S_2^{\mathrm{T}}.$$

1.10 证明引理1.4.11.

1.11 设$X\in\mathbb{R}^m$, $Y\in\mathbb{R}^n$, $Z\in\mathbb{R}^p$, 它们都是列向量. 证明

$$Y\otimes X\otimes Z=\left(W_{[m,n]}\otimes I_p\right)X\otimes Y\otimes Z;$$

$$X\otimes Z\otimes Y=\left(I_m\otimes W_{[n,p]}\right)X\otimes Y\otimes Z;$$

$$Z\otimes Y\otimes X=\left(W_{[n,p]}\otimes I_m\right)\left(I_n\otimes W_{[m,p]}\right)\left(W_{[m,n]}\otimes I_p\right)X\otimes Y\otimes Z.$$

1.12 证明换位矩阵的两种表示形式: (1.5.9) 和(1.5.10).

1.13 设A 为$m\times n$ 矩阵, 证明公式

$$V_c(A^{\mathrm{T}})=W_{[n,m]}V_c(A);$$

$$V_r(A^{\mathrm{T}})=W_{[m,n]}V_r(A).$$

1.14 找出例1.3.4 (囚徒的困境)的Nash 均衡点.

第 2 章　矩阵的左半张量积

本章我们首先回顾几种常用矩阵乘法: 普通矩阵乘法、张量积以及Hadamard积的一些基本性质, 从而理解作为矩阵乘积的一些基本要求. 然后我们介绍和分析一种关于三维数组的处理方法: 立方阵和它的乘法. 通过比较和分析, 我们提出一种称为左半张量积的矩阵乘法, 并介绍它的一些基本性质. 最后将它的一些直接应用作为例子来说明这种新矩阵乘法的意义和效果.

2.1　矩阵乘法的一些基本性质

首先, 让我们回忆一下几种常用的矩阵乘法: 普通矩阵乘法、张量积 (也称Kronecker 积) [65] 和Hadamard 积[124], 看看矩阵乘法要求的最基本要素是什么.

设$A=(a_{ij})\in M_{m\times n}$, $B=(b_{ij})\in M_{s\times t}$, 那么

$$A\otimes B=\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B\\ \vdots & \vdots & & \vdots\\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \tag{2.1.1}$$

称为矩阵A 和B 的张量积.

设$A=(a_{ij})\in M_{m\times n}$, $B=(b_{ij})\in M_{m\times n}$, 那么

$$A\odot B=\begin{bmatrix} a_{11}b_{11} & a_{12}b_{12} & \cdots & a_{1n}b_{1n}\\ \vdots & \vdots & & \vdots\\ a_{m1}b_{m1} & a_{m2}b_{m2} & \cdots & a_{mn}b_{mn} \end{bmatrix} \tag{2.1.2}$$

称为矩阵A和B的 Hadamard 积.

普通矩阵乘法、张量积和Hadamard 积这三种矩阵乘法都满足两个最基本的规则:

(1) 结合律

$$\begin{aligned}&(AB)C=A(BC),\\&(A\otimes B)\otimes C=A\otimes (B\otimes C),\\&(A\odot B)\odot C=A\odot (B\odot C).\end{aligned}$$

(2) 分配律

$$(A+B)C = AC + BC, \quad A(B+C) = AB + AC;$$
$$(A+B)\otimes C = A\otimes C + B\otimes C, \quad A\otimes(B+C) = A\otimes B + A\otimes C;$$
$$(A+B)\odot C = A\odot C + B\odot C, \quad A\odot(B+C) = A\odot B + A\odot C.$$

这两条性质使具有合适维数的矩阵集合在矩阵乘法(普通矩阵乘法、张量积和Hadamard 积) 下构成一种代数结构(结合代数). 这两条应当是作为矩阵乘法的基本要求.

为了讨论的方便, 下面我们不加证明地介绍张量积和Hadamard 积的一些基本性质, 这些性质将在后面用到.

命题 2.1.1 *设$A, B, C, \cdots$ 为矩阵, $a \in \mathbb{R}$, 那么*

(1)

$$(aA)\otimes B = a(A\otimes B) = A\otimes(aB), \quad a\in\mathbb{R}. \tag{2.1.3}$$

$$(A\otimes B)^{\mathrm{T}} = A^{\mathrm{T}}\otimes B^{\mathrm{T}}. \tag{2.1.4}$$

(2) *如果$A = A^{\mathrm{T}}$, $B = B^{\mathrm{T}}$, 那么*

$$(A\otimes B)^{\mathrm{T}} = A\otimes B. \tag{2.1.5}$$

(3) *如果A 和B 都是方阵, 那么*

$$\operatorname{tr}(A\otimes B) = \operatorname{tr}(A)\operatorname{tr}(B). \tag{2.1.6}$$

(4) *如果A 和B 都可逆, 那么*

$$(A\otimes B)^{-1} = A^{-1}\otimes B^{-1}. \tag{2.1.7}$$

(5) *对于两个向量X 和Y, 我们有*

$$XY^{\mathrm{T}} = Y^{\mathrm{T}}\otimes X = X\otimes Y^{\mathrm{T}}. \tag{2.1.8}$$

(6) *设X 和Y 是矩阵A 和B 分别属于特征值μ 和λ 的特征向量, 则$\mu\lambda$ 是$A\otimes B$ 的特征值, 并且$X\otimes Y$ 是$A\otimes B$ 的属于特征值$\mu\lambda$ 的特征向量.*

(7) *设$\mu_1, \cdots, \mu_m$ 是A 的特征值, $\lambda_1, \cdots, \lambda_n$ 是B 的特征值, 则$A\otimes B$ 的特征值是$\{\mu_i\lambda_j \mid i = 1, \cdots, m,\ j = 1, \cdots, n\}$.*

对于对称矩阵A, 记号$A > 0$, $A \geq 0$, $A < 0$ 和$A \leq 0$ 分别表示A 正定, 半正定, 负定和半负定.

Hadamard 积有下面的重要性质:

命题 2.1.2 (1) 设$A \in M_n$, $E = (\underbrace{1\ 1\ \cdots\ 1}_{n})^{\mathrm{T}}$ (或简记$E = \mathbf{1}_n$), 则

$$A \odot (EE^{\mathrm{T}}) = A = (EE^{\mathrm{T}}) \odot A. \tag{2.1.9}$$

(2) 对于任意两个矩阵$A, B \in M_{m\times n}$ 有

$$A \odot B = B \odot A. \tag{2.1.10}$$

(3) 设$X, Y \in \mathbb{R}^n$ 是两个向量, 则

$$(XX^{\mathrm{T}}) \odot (YY^{\mathrm{T}}) = (X \odot Y)(X \odot Y)^{\mathrm{T}}. \tag{2.1.11}$$

(4) (Schur 定理) 设$A, B \in M_n$ 是对称矩阵.
如果$A \geq 0, B \geq 0$, 则$A \circ B \geq 0$;
如果$A > 0, B > 0$, 则$A \circ B > 0$.
(5) (Oppenbeim 定理) 设$A, B \in M_n$ 是对称矩阵. 如果$A \geq 0$, $B \geq 0$, 则

$$\det(A \odot B) \geq \det(A)\det(B). \tag{2.1.12}$$

普通矩阵乘法和张量积之间的一个最基本的关系为:

命题 2.1.3 设A, B, C, D 都有合适的维数, 那么

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD). \tag{2.1.13}$$

实际上, 在第1章中就曾经用到矩阵的张量积的一些性质, 特别是命题2.1.3 中的性质. 这个性质在以后也会经常用到.

设δ_i^n 是I_n 的第i 列, 并且定义一个$n^2 \times n$ 矩阵

$$H_n = \mathrm{diag}(\delta_1^n, \cdots, \delta_n^n) = \begin{bmatrix} \delta_1^n & 0 & \cdots & 0 \\ 0 & \delta_2^n & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \delta_n^n \end{bmatrix}.$$

那么我们有

命题 2.1.4 设$A, B \in M_{m\times n}$, 那么

$$A \odot B = H_m^{\mathrm{T}}(A \otimes B)H_n. \tag{2.1.14}$$

2.2 立方阵

上一章曾经提到, 为了处理3 维数组, 一些学者引入了立方阵的概念. 它确实也得到一些较好的实际应用. 本节将对此作一个简单介绍.

定义 2.2.1 称形如

$$X=\{x_{i,j,k}\,|\,i=1,\cdots,n;\ j=1,\cdots,p;\ k=1,\cdots,q\} \tag{2.2.1}$$

的3 维数组为立方阵.

在上述定义中, 3 维数组被排列成立方的形式. 如果一个3 维数组被排列成立方的形式, 那么一个根本问题就是如何表示和计算它们? 实际上, 按照其运算的定义不难看出, 归根到底, 它们还是相当于排列成矩阵的形式.

定义 2.2.2 称(2.2.1) 中的立方阵X 被排列成$\binom{1}{23}$ 形式(或1-行形式), 记为X_{23}^{1}, 如果

$$X_{23}^{1}=\begin{bmatrix} x_{111} & \cdots & x_{11q} & \cdots & x_{1p1} & \cdots & x_{1pq} \\ x_{211} & \cdots & x_{21q} & \cdots & x_{2p1} & \cdots & x_{2pq} \\ \vdots & & \vdots & & \vdots & & \vdots \\ x_{n11} & \cdots & x_{n1q} & \cdots & x_{np1} & \cdots & x_{npq} \end{bmatrix}.$$

同理, X 也可以被排列成其他11 种矩阵形式: X_{32}^{1}, X_{13}^{2}, X_{31}^{2}, X_{12}^{3}, X_{21}^{3}, X_{3}^{12}, X_{3}^{21}, X_{2}^{13}, X_{2}^{31}, X_{1}^{23}和X_{1}^{32}. 这里不包括行向量和列向量形式.

通过块转置可以很容易得到这些矩阵表示之间的关系.

命题 2.2.3 一个立方阵的不同矩阵表示之间满足下列关系:

$$(X_{23}^{1})^{\mathrm{T}(1\times q)}=X_{13}^{2},(X_{12}^{3})^{\mathrm{T}(1\times p)}=X_{32}^{1},(X_{21}^{3})^{\mathrm{T}(1\times n)}=X_{31}^{2},$$

$$(X_{3}^{12})^{\mathrm{T}(p\times 1)}=X_{1}^{32},(X_{2}^{13})^{\mathrm{T}(q\times 1)}=X_{1}^{23},(X_{3}^{21})^{\mathrm{T}(n\times 1)}=X_{2}^{31},$$

$$(X_{23}^{1})^{\mathrm{T}(n\times q)}=X_{3}^{21},(X_{13}^{2})^{\mathrm{T}(p\times q)}=X_{3}^{12},(X_{32}^{1})^{\mathrm{T}(n\times p)}=X_{2}^{31},$$

$$(X_{1}^{32})^{\mathrm{T}(p\times n)}=X_{31}^{2},(X_{2}^{13})^{\mathrm{T}(q\times p)}=X_{12}^{3},(X_{1}^{23})^{\mathrm{T}(q\times n)}=X_{21}^{3}.$$

在文献[5], [6], [97] 中定义了立方阵和普通矩阵之间的两种乘法.

定义 2.2.4 设X 是一个$n\times p\times q$ 立方阵.

(1) A 是一个$m\times p$ 矩阵. 定义$Y=A\oslash X$, 其中Y 是一个$n\times m\times q$ 立方阵, 并且满足$y_{ijk}=\sum\limits_{s=1}^{p}a_{js}x_{isk}$.

(2) A 是一个$q \times m$ 矩阵. 定义$Y = X \oslash A$, 其中Y 是一个$n \times p \times m$ 立方阵, 并且满足$y_{ijk} = \sum\limits_{s=1}^{q} x_{ijs}a_{sk}$.

(3) A 是一个$m \times n$ 矩阵. 定义$Y = [A][X]$, 其中Y 是一个$m \times p \times q$ 立方阵, 并且满足$y_{ijk} = \sum\limits_{s=1}^{n} a_{is}x_{sjk}$.

(4) A 是一个$n \times m$ 矩阵. 定义$Y = [X][A]$, 其中Y 是一个$m \times p \times q$ 立方阵, 并且满足$y_{ijk} = \sum\limits_{s=1}^{n} x_{sjk}a_{si}$.

注

(1)在文献[5] 等中, 对于乘法$\oslash$, 乘法记号是省略的, 即$AX := A \oslash X$. 我们使用记号$\oslash$ 是为了避免混淆.

(2)当X 被排列成$\binom{12}{3}$ 形式时, 即将X 表示成

$$X = \begin{bmatrix} X^1 \\ \vdots \\ X^n \end{bmatrix}, \quad \text{这里} \quad X^k = \begin{bmatrix} x_{k11} & \cdots & x_{k1q} \\ \vdots & & \vdots \\ x_{kp1} & \cdots & x_{kpq} \end{bmatrix}.$$

那么

$$A \oslash X = \begin{bmatrix} AX^1 \\ \vdots \\ AX^n \end{bmatrix}, \quad X \oslash A = \begin{bmatrix} X^1A \\ \vdots \\ X^nA \end{bmatrix}.$$

(3)设X 如(2) 中一样排列成$\binom{12}{3}$ 形式, 并且$p = q$, 可定义X 的迹为

$$\operatorname{tr}(X) = \begin{bmatrix} \operatorname{tr}(X^1) \\ \vdots \\ \operatorname{tr}(X^n) \end{bmatrix}. \tag{2.2.2}$$

(4)设$p = q$, 且$b \in \mathbb{R}^p$, 定义二次型$b^{\mathrm{T}}Xb$ 为

$$b^{\mathrm{T}}Xb = \begin{bmatrix} b^{\mathrm{T}}X^1b \\ \vdots \\ b^{\mathrm{T}}X^nb \end{bmatrix}. \tag{2.2.3}$$

在文献[107] 中证明了立方阵乘法的一些基本性质.

命题 2.2.5 [107] 设A, B, C 是普通矩阵, X, Y 是立方阵, α, β 是实数. 假设它们具有使下面的式子有定义的合适维数, 那么我们有

(1) $[I][X]=X$;

(2) $[\alpha A+\beta B][X]=\alpha[A][X]+\beta[B][X]$;

(3) $[A][\alpha X+\beta Y]=\alpha[A][X]+\beta[A][Y]$;

(4) $[A][X]=[X][A^{\mathrm{T}}]$;

(5) $[A][(L\oslash X)M]=(L\oslash([A][X]))\,M$;

(6) $[AB][X]=[A][[B][X]]$.

实际上, 在普通矩阵和立方阵之间还有其他几种乘法, 这在文献[125] 中有详细叙述.

定义 2.2.6 设X 是一个$n\times p\times q$ 立方阵.

(1) A 是一个$m\times n$ 矩阵. 定义$Y=A\prod_{(1)}X$, 其中Y 是一个$m\times p\times q$ 立方阵, 并且满足$y_{ijk}=\sum\limits_{s=1}^{n}a_{is}x_{sjk}$.

(2) A 是一个$m\times p$ 矩阵. 定义$Y=A\prod_{(2)}X$, 其中Y 是一个$n\times m\times q$ 立方阵, 并且满足$y_{ijk}=\sum\limits_{s=1}^{p}a_{js}x_{isk}$.

(3) A 是一个$m\times q$ 矩阵. 定义$Y=A\prod_{(3)}X$, 其中Y 是一个$n\times p\times m$ 立方阵, 并且满足$y_{ijk}=\sum\limits_{s=1}^{q}a_{ks}x_{ijs}$.

(4) A 是一个$n\times m$ 矩阵. 定义$Y=X\prod_{(1)}A$, 其中Y 是一个$m\times p\times q$ 立方阵, 并且满足$y_{ijk}=\sum\limits_{s=1}^{n}x_{sjk}a_{si}$.

(5) A 是一个$p\times m$ 矩阵. 定义$Y=X\prod_{(2)}A$, 其中Y 是一个$n\times m\times q$ 立方阵, 并且满足$y_{ijk}=\sum\limits_{s=1}^{p}x_{isk}a_{sj}$.

(6) A 是一个$q\times m$ 矩阵. 定义$Y=X\prod_{(3)}A$, 其中Y 是一个$n\times p\times m$ 立方阵, 并且满足$y_{ijk}=\sum\limits_{s=1}^{q}x_{ijs}a_{sk}$.

注 $\oslash$ 乘法和括号乘法都是这六种立方阵乘法的特殊情形. 实际上有

$$A\oslash X=A\prod_{(2)}X,\quad X\oslash A=X\prod_{(3)}A;$$

$$[A][X]=A\prod_{(1)}X,\quad [X][A]=X\prod_{(1)}A.$$

另外, 上面的立方阵乘法可以在合适的矩阵表示下转化为普通矩阵乘法.

命题 2.2.7 设X 是一个$n\times p\times q$ 立方阵, A 是一个普通矩阵且具有使下面的式子有定义的合适维数. 那么在下列情况下立方阵乘法可以转化为普通矩阵乘法.

(1) 如果X 排列成X_{23}^1 或X_{32}^1, 且$Y=A\prod_{(1)}X$, 那么

$$Y_{23}^1=AX_{23}^1,\quad Y_{32}^1=AX_{32}^1.$$

(2) 如果X 排列成X_{13}^2 或X_{31}^2, 且$Y=A\prod_{(2)}X$, 那么

$$Y_{13}^2=AX_{13}^2,\quad Y_{31}^2=AX_{31}^2.$$

(3) 如果X 排列成X_{12}^3 或X_{21}^3, 且$Y=A\prod_{(3)}X$, 那么

$$Y_{12}^3=AX_{12}^3,\quad Y_{21}^3=AX_{21}^3.$$

(4) 如果X 排列成X_1^{23} 或X_1^{32}, 且$Y=X\prod_{(1)}A$, 那么

$$Y_1^{23}=X_1^{23}A,\quad Y_1^{32}=X_1^{32}A.$$

(5) 如果X 排列成X_2^{13} 或X_2^{31}, 且$Y=X\prod_{(2)}A$, 那么

$$Y_2^{13}=X_2^{13}A,\quad Y_2^{31}=X_2^{31}A.$$

(6) 如果X 排列成X_{12}^3 或X_{21}^3, 且$Y=X\prod_{(3)}A$, 那么

$$Y_{12}^3=X_{12}^3A,\quad Y_{21}^3=X_{21}^3A.$$

虽然立方阵和普通矩阵之间的$\oslash$ 乘法和括号乘法在统计学中已经得到一些成功的应用, 但是它们的缺点也是显而易见的, 例如: ①不满足结合律, 这将会给我们的使用带来很大不便;②定义比较复杂并且不能推广到更高维数的情形; ③与普通矩阵乘法不相容; ④正如定义2.2.6 中介绍的其他乘法, 研究这两种乘法也没有什么特殊意义. 既然定义这么多乘法很麻烦, 因此有必要引入一种通用的关于高维数组的乘法.

2.3 左半张量积

从上一节可以看出, 即使只考虑3 维数组(立方阵), 也可以定义许多不同的表示形式和乘法, 而且很难甚至不可能将它们推广到高维情形. 另外, 这些乘法与

普通矩阵乘法不相容(即混合乘法不满足结合律)，所以应用起来很不方便. 因此需要定义一种统一的高维数组乘法，并且它与普通矩阵乘法相容. 左半张量积就是基于这个目的而提出来的.

我们已经提到过，在计算机存储中并不是将数据排列成二维形式来实现矩阵运算，而是使用“指针”、“指针的指针”、“指针的指针的指针”等. 实际上，这种排列方式比矩阵形式更好，因为通过指针的层次，我们不仅可以将数据排列成矩阵形式，而且还可以将数据排列成立方甚至是更高维的形式. 现在的问题是：我们能否创造一种新的运算，使它能够进行这种多层次数组的计算？为此我们引入左半张量积的概念. 首先让我们回顾一下普通的矩阵乘法.

例 2.3.1　设U 和V 分别是m 和n 维向量空间，且$F \in L(U \times V, \mathbb{R})$，即$F$ 是一个从$U \times V$ 到$\mathbb{R}$ 的双线性映射. 设$\{u_1, \cdots, u_m\}$ 和$\{v_1, \cdots, v_n\}$ 分别是U 和V 的基底. 我们称$S = (s_{ij})$ 是F 的结构常数，这里

$$s_{ij} = F(u_i, v_j), \quad i = 1, \cdots, m, \quad j = 1, \cdots, n.$$

设$X = \sum\limits_{i=1}^{m} a_i u_i \in U$，简记为$X = (a_1, \cdots, a_m)^{\mathrm{T}} \in U$，$Y = \sum\limits_{i=1}^{n} b_i v_i \in V$ 或$Y = (b_1, \cdots, b_n)^{\mathrm{T}} \in V$，那么

$$F(X, Y) = X^{\mathrm{T}} S Y. \tag{2.3.1}$$

下面我们给出上面式子的一个可行计算过程：记S 的行为$S^1, \cdots, S^n$，那么可以看到F 可以通过下面两步计算得到

第 1 步：计算$x_1 S^1$，$x_2 S^2$，$\cdots$，$x_m S^m$，再将它们全部加起来；

第 2 步：向量$\sum\limits_{i=1}^{m} x_i S^i$ 乘以Y (实际上就是做标准内积). □

这个计算过程促使我们引入一种新的矩阵乘法，定义如下：

定义 2.3.2　设T 是一个np 维行向量，X 是一个p 维列向量. 将T 分割成p 个等长的块$T^1, \cdots, T^p$，它们每一个都是$1 \times n$ 维行向量. 定义左半张量积$\ltimes$ 为

$$T \ltimes X = \sum_{i=1}^{p} T^i x_i \in \mathbb{R}^n. \tag{2.3.2}$$

我们利用这种新乘法继续上面的例2.3.1，并且给出一个可行的算法.

例 2.3.3 (继续例2.3.1)　我们将F 的结构常数重新排列成

$$T := (s_{11}, \cdots, s_{1n}, \ \cdots, s_{m1}, \cdots, s_{mn}),$$

称之为结构矩阵, 它是一个mn 维的行向量, 由指标(i,j) 按照索引$Id(i,j;m,n)$ 排列. 下面的公式给出了与(2.3.1) 一样的结果.

$$F(X,Y) = T \ltimes X \ltimes Y. \tag{2.3.3}$$

□

可以通过直接计算检验(2.3.3) 的正确性. 那么(2.3.3) 的好处在哪呢? 注意到(2.3.3) 使用1 维向量实现了2 维数组的运算. 既然可以用1维向量进行2 维数组的运算, 那么我们就希望利用1 维向量实现有限维数组的运算. 如果可以这样做的话, (2.3.3) 就比(2.3.1) 更有用. 让我们接着看下面的例子.

例 2.3.4 设U, V 和W 分别是m, n 和t 维向量空间, 且$F \in L(U \times V \times W, \mathbb{R})$. 设$\{u_1,\cdots,u_m\}$, $\{v_1,\cdots,v_n\}$, $\{w_1,\cdots,w_t\}$ 分别是U, V, W 的基底. 我们定义结构常数为

$$s_{ijk} = F(u_i,v_j,w_k), \quad i=1,\cdots,m, \quad j=1,\cdots,n, \quad k=1,\cdots,t.$$

现在我们构造F 的结构矩阵S, 它是一个由$\{i,j,k\}$ 按照索引$Id(i,j,k;m,n,t)$ 排列成的mnt 维行向量, 即

$$S = (s_{111},\cdots,s_{11t},\cdots,s_{1n1},\cdots,s_{1nt},\cdots,s_{mn1},\cdots,s_{mnt}).$$

那么对于$X \in U$, $Y \in V$, $Z \in W$, 很容易验证

$$F(X,Y,Z) = S \ltimes X \ltimes Y \ltimes Z.$$

□

我们可以看到在半张量乘法中, 算子$\ltimes$ 可以自动找到“指针”的不同的层, 再进行相应的计算.

显然, 例2.3.4 中的结构和计算可以推广到在任意有限个向量空间上定义的多线性映射.

现在我们可以给出矩阵的左半张量积的一般定义了.

定义 2.3.5 (1) 设$X=(x_1,\cdots,x_s)$ 是一个行向量, $Y=(y_1,\cdots,y_t)^{\mathrm{T}}$ 是一个列向量.

情形1：如果t 是s 的因子, 即$s=t\times n$, 那么n 维行向量

$$\langle X,Y\rangle_L := \sum_{k=1}^{t} X^k y_k \in \mathbb{R}^n \tag{2.3.4}$$

称为是X 和Y 的左半张量内积, 这里$X=(X^1,\cdots,X^t)$, $X^i\in\mathbb{R}^n$, $i=1,\cdots,t$.

情形2 : 如果s 是t 的因子, 即$t=s\times n$, 那么n 维列向量

$$\langle X,Y\rangle_L:=(\langle Y^{\mathrm{T}},X^{\mathrm{T}}\rangle_L)^{\mathrm{T}}\in\mathbb{R}^n \tag{2.3.5}$$

也称为是X 和Y 的左半张量内积.

(2) 设$M\in M_{m\times n}$, $N\in M_{p\times q}$. 如果n 是p 的因子或者p 是n 的因子, 称$C=M\ltimes N$ 是M 和N 的左半张量积, 如果C 由$m\times q$ 个块组成, 即$C=(C^{ij})$, 并且

$$C^{ij}=\langle M^i,N_j\rangle_L,\quad i=1,\cdots,m,\quad j=1,\cdots,q.$$

注 定义2.3.5 的第(1) 条中, 如果$t=s$, 那么左半张量内积就变成标准内积. 因而在定义2.3.5 的第(2) 条中, 如果$n=p$, 那么矩阵的左半张量积就退化成普通矩阵乘法. 因此说左半张量积是普通矩阵乘法的推广.

给定两个矩阵$M_{m\times n}$ 和$N_{p\times q}$, $M\ltimes N$ 有定义当且仅当下列两种情况之一成立:

(1) 如果$n\%p=0$, 那么此时乘积$M\ltimes N$ 的维数是$m\times\frac{nq}{p}$;

(2) 如果$p\%n=0$, 那么此时乘积$M\ltimes N$的维数是$\frac{mp}{n}\times q$.

注 (1) 为方便计, 当$n=p$ 时, 我们称M, N 满足等维数条件, 当$n\%p=0$ 或$p\%n=0$ 时, 我们称M, N 满足倍维数条件. 半张量积就是将矩阵乘法从满足等维数条件的矩阵对推广到满足倍维数条件的矩阵对.

(2) 在本书第4 章第2 节, 我们将定义两个具有任意维数的矩阵的左(右)半张量积. 但它仅限于该节, 在本书其余部分, 我们仅对满足倍维数条件的矩阵对定义矩阵的左(右)半张量积.

例 2.3.6 (1) 设$X=\begin{bmatrix}1&2&-1&3\end{bmatrix}$, $Y=\begin{bmatrix}2&1\end{bmatrix}^{\mathrm{T}}$, 那么

$$\langle X,Y\rangle_L=\begin{bmatrix}1&2\end{bmatrix}\times 2+\begin{bmatrix}-1&3\end{bmatrix}\times 1=\begin{bmatrix}1&7\end{bmatrix}.$$

(2) 设$X=(2\ 1)$, $Y=(1\ 2\ -1\ 3)^{\mathrm{T}}$, 那么

$$\langle X,Y\rangle_L=2\times\begin{bmatrix}1\\2\end{bmatrix}+1\times\begin{bmatrix}-1\\3\end{bmatrix}=\begin{bmatrix}1\\7\end{bmatrix}.$$

(3) 设

$$X=\begin{bmatrix}1&2&-1&2\\0&1&2&3\\3&3&1&1\end{bmatrix},\quad Y=\begin{bmatrix}1&2\\-1&3\end{bmatrix}.$$

那么

$$X \ltimes Y = \begin{bmatrix} (1\,2)-(-1\,2) & 2(1\,2)+3(-1\,2) \\ (0\,1)-(2\,3) & 2(0\,1)+3(2\,3) \\ (3\,3)-(1\,1) & 2(3\,3)+3(1\,1) \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 & -1 & 10 \\ -2 & -2 & 6 & 11 \\ 2 & 2 & 9 & 9 \end{bmatrix}.$$

□

注

(1) 两个矩阵的左半张量积的维数可以很容易根据前一个矩阵的第二个指标和后一个矩阵的第一个指标的公因子消去来计算得到. 例如:

$$A_{p\times qr} \ltimes B_{r\times s} \ltimes C_{qst\times l} = (A \ltimes B)_{p\times qs} \ltimes C_{qst\times l} = (A \ltimes B \ltimes C)_{pt\times l}.$$

在第一个等式中消去了r, 第二个等式中消去了qs. 这显然可以看作是普通矩阵乘法的推广, 因为对于普通矩阵乘法$A_{p\times s}B_{s\times q} = (AB)_{p\times q}$, 可以认为是消去了$s$.

(2)不同于普通的矩阵乘法, 对于半张量积, 即使$A \ltimes B$ 和$B \ltimes C$ 都有定义, 但是$A \ltimes B \ltimes C = (A \ltimes B) \ltimes C$ 也可能没有定义. 例如, $A \in M_{3\times 4}$, $B \in M_{2\times 3}$, $C \in M_{9\times 1}$.

(3)在左半张量积$\ltimes$ 下, 对于任意矩阵A 和$r \in \mathbb{R}$ 都有

$$r \ltimes A = A \ltimes r = rA.$$

这是因为r 可以被看作是一个1×1 矩阵, 因此在左半张量积下我们不需要区分数乘和矩阵乘法. 显然对于左半张量积, 1 是"通用单位元", 0 是"通用零元".

通过对矩阵的普通乘法、张量积和左半张量积的比较发现, 它们之间的一个主要差别是: 普通乘法是1 对1 型的乘法, 就是一个矩阵的一个元素和另一个矩阵的一个元素相乘; 张量积是1对矩阵型的乘法; 半张量积是1 对块型的乘法. 似乎左半张量积是介于两者之间的, 因此我们将它称为半张量积.

下面的例子说明, 普通矩阵乘法在一些"合法"运算之后会出现"非法"项, 这给美妙的矩阵运算带来了一丝阴云. 而当标准矩阵乘法推广到左半张量积时, 一切又都变得自然而合理了. 这也许可以给出推广普通矩阵乘法的一个有力支持.

例 2.3.7 设$X,Y,Z,W\in\mathbb{R}^n$ 都是列向量. 因为$Y^{\mathrm{T}}Z$ 是一个数, 于是

$$(XY^{\mathrm{T}})(ZW^{\mathrm{T}})=X(Y^{\mathrm{T}}Z)W^{\mathrm{T}}=(Y^{\mathrm{T}}Z)(XW^{\mathrm{T}})\in M_{n\times n}. \tag{2.3.6}$$

如果继续利用矩阵乘法的结合律应该有

$$(Y^{\mathrm{T}}Z)(XW^{\mathrm{T}})=Y^{\mathrm{T}}(ZX)W^{\mathrm{T}}. \tag{2.3.7}$$

现在我们就有一个疑问: ZX 是什么? 这说明普通的矩阵乘法并不是那么相容的.

因为我们可以将普通的矩阵乘法看成左半张量积的特殊情形, 于是

$$(XY^{\mathrm{T}})(ZW^{\mathrm{T}})=Y^{\mathrm{T}}\ltimes(Z\ltimes X)\ltimes W^{\mathrm{T}}. \tag{2.3.8}$$

可以证明, (2.3.8) 确实成立. 因此, 当将普通矩阵乘法推广到一般的左半张量积时, 上述的不相容性就消失了. □

下面的两个简单应用将进一步说明引入半张量积的合理性.

例 2.3.8 回顾例1.3.1. 设X 和Y 是李代数 $\mathcal{G}$ 中的两个向量, 它们可以表示成 $X=\sum\limits_{i=1}^{n}x_iE_i$, $Y=\sum\limits_{i=1}^{n}y_iE_i$, 或简记为$X=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$, $Y=(y_1,y_2,\cdots,Y_n)^{\mathrm{T}}$. 如果我们用$M_g=M_2^{13}$ 作为结构矩阵, 那么直接计算有

$$[X,Y]=X^{\mathrm{T}}\ltimes M_g\ltimes Y. \tag{2.3.9}$$

同样, 如果我们用$M^g=M_{12}^3$ 作为结构矩阵, 那么

$$[X,Y]=M^g\ltimes X\ltimes Y. \tag{2.3.10}$$

□

例 2.3.9 再回顾例1.3.2. 对于$X_i=\sum\limits_{k=1}^{n}x_k^id_k, i=1,2,\cdots,r$ 和$Y_j=\sum\limits_{k=1}^{n}y_k^je_k$, $j=1,2,\cdots,s$, 将它们简记为$X_i=(x_1^i,x_2^i,\cdots,x_n^i)^{\mathrm{T}}$ 和$\sigma_j=(y_1^j,y_2^j,\cdots,y_n^j)$. 利用(1.3.8), 很容易验证

$$\begin{aligned}&\omega(X_1,X_2,\cdots,X_r;\sigma_1,\sigma_2,\cdots,\sigma_s)\\&=\sigma_s\ltimes\sigma_{s-1}\ltimes\cdots\ltimes\sigma_1\ltimes M_\omega\ltimes X_1\ltimes X_2\ltimes\cdots\ltimes X_r.\end{aligned} \tag{2.3.11}$$

特别地, 给定一个向量$X\in V$, 映射$i_X: T_s^r\to T_s^{r-1}$ 定义为[10]

$$i_X(\omega)=\omega(X,\cdots;\cdots). \tag{2.3.12}$$

利用左半张量积, $i_X(\omega)$ 的张量向量可以简单表示成

$$M_{i_X(\omega)} = M_\omega \ltimes X.$$

同理, 给定一个余向量$\sigma \in V^*$, 映射$i_\sigma:\ T_s^r \to T_{s-1}^r$ 定义为

$$i_\sigma(\omega) = \omega(\cdots;\sigma,\cdots). \tag{2.3.13}$$

那么$i_\sigma(\omega)$ 的张量向量可以表示成

$$M_{i_\sigma(\omega)} = \sigma \ltimes M_\omega.$$

□

我们可以将上例中的自变量推广到不同维数向量空间乘积空间的一般情况: 设η 是一个多线性映射, 即

$$\eta \in L\left(V_1 \times V_2 \times \cdots \times V_k; W\right),$$

其中$\dim(V_i) = n_i$, $\dim(W) = n_0$. 给定$X \in V_t$, 我们定义

$$i_X^t(\eta): L\left(V_1 \times \cdots \times V_k; W\right) \to L\left(V_1 \times \cdots V_{t-1} \times V_{t+1} \times \cdots \times V_k; W\right),$$

为

$$\begin{gathered} i_X^t(\eta)(Y_1,\cdots,Y_{k-1}) = \eta(Y_1,\cdots,Y_{t-1},X,Y_t,\cdots,Y_{k-1}), \\ \forall Y_i \in V_i,\ i<t; Y_i \in V_{i+1},\ i \geq t. \end{gathered}$$

不难证明(这里留作练习)

命题 2.3.10 设M_η 和M_ζ 分别是η 和$\zeta = i_X^t(\eta)$ 的结构矩阵, 那么

$$M_\zeta = M_\eta \ltimes (I_{n_1+\cdots+n_{t-1}} \otimes X). \tag{2.3.14}$$

接下来我们定义矩阵的幂.

定义 2.3.11 给定一个矩阵$A \in M_{p\times q}$. 假设$p\%q = 0$ 或$q\%p = 0$, 我们递推地定义A^n, $n > 0$ 为

$$\begin{cases} A^1 = A, \\ A^{k+1} = A^k \ltimes A, \quad k = 1,2,\cdots. \end{cases}$$

注 容易验证矩阵的幂是有定义的. 并且, 如果$p = sq$, 其中s 是一个正整数, 那么A^k 的维数是$(s^k q, q)$; 如果$q = sp$, 那么A^k 的维数是$(p, s^k p)$.

例 2.3.12 (1) 如果X 是一个行向量或列向量, 那么根据定义2.3.11, X^n 总是有定义的. 实际上, 容易看出当X, Y 都是列向量时有

$$X \ltimes Y = X \otimes Y, \tag{2.3.15}$$

当X, Y 都是行向量时有

$$X \ltimes Y = Y \otimes X. \tag{2.3.16}$$

因此不管是上面哪种情况都有

$$X^k = \underbrace{X \otimes \cdots \otimes X}_{k}. \tag{2.3.17}$$

(2) 设$X \in \mathbb{R}^m$, $Y \in \mathbb{R}^p$ 都是列向量, 给定两个矩阵$A \in M_{m\times n}$, $B \in M_{p\times q}$, 那么

$$(AX) \ltimes (BY) = (A \otimes B)(X \ltimes Y). \tag{2.3.18}$$

因此

$$(AX)^k = (\underbrace{A \otimes \cdots \otimes A}_{k})X^k. \tag{2.3.19}$$

(3) 设X, Y 是行向量, A, B 是两个有合适维数的矩阵. 那么

$$(XA) \ltimes (YB) = (X \ltimes Y)(B \otimes A). \tag{2.3.20}$$

因此

$$(XA)^k = X^k(\underbrace{A \otimes \cdots \otimes A}_{k}). \tag{2.3.21}$$

(4) 考虑变元为$x \in \mathbb{R}^n$ 的k 次齐次多项式全体的集合, 记为B_n^k. 在普通的多项式加法和数乘下, B_n^k 是一个向量空间. 显然x^k 包含一组基底(但它不是基底, 因为有冗余的元素). 因此, 任意的$p(x) \in B_n^k$ 都可以表示成$p(x) = Cx^k$, 但是系数$C \in \mathbb{R}^{n^k}$ 不唯一. □

在上面的证明中我们将一些细节的验证留给了读者.

定义 2.3.13 给定$A_{m\times n}$ 和$B_{p\times q}$.

(1) 如果$n\%p = 0$, 我们就记$A \succ B$; 如果$p\%n = 0$, 记为$A \prec B$. 如果$A \ltimes B$ 有定义, 那么就有$A \succ B$ 或$B \prec A$.

(2) 如果$A \succ B$ 且$n = tp$, 为了强调t 我们记为$A \succ_t B$; 相反, 如果$A \prec B$ 且$nt = p$, 我们记为$A \prec_t B$.

(3) 对于一个有限的矩阵序列F: $A_1, \cdots, A_n$, 如果有$A_1 \succ A_2 \succ \cdots \succ A_n$ 或者$A_1 \prec A_2 \prec \cdots \prec A_n$, 那么称这个序列为一个有序链.

命题 2.3.14 (1) 给定$A_{m\times n}$ 和$B_{p\times q}$, 它们有合适的维数. 那么

$$
\begin{aligned}
A \ltimes B &= \begin{bmatrix} A_1 \ltimes B^1 & \cdots & A_1 \ltimes B^q \\ \vdots & & \vdots \\ A_m \ltimes B^1 & \cdots & A_m \ltimes B^q \end{bmatrix} = \begin{bmatrix} A_1 \ltimes B \\ \vdots \\ A_m \ltimes B \end{bmatrix} \\
&= \begin{bmatrix} A \ltimes B^1 & \cdots & A \ltimes B^q \end{bmatrix},
\end{aligned}
\tag{2.3.22}
$$

其中$A_i (i = 1, \cdots, m)$ 表示A 的第i 行, $B^j (j = 1, \cdots, k)$ 表示B 的第j 列.

(2) 假设$A \succ_t B$ (或$A \prec_t B$). 将A 和B 分割成如下的分块形式

$$
A = \begin{bmatrix} A^{11} & \cdots & A^{1s} \\ \vdots & & \vdots \\ A^{r1} & \cdots & A^{rs} \end{bmatrix}, \quad B = \begin{bmatrix} B^{11} & \cdots & B^{1t} \\ \vdots & & \vdots \\ B^{s1} & \cdots & B^{st} \end{bmatrix}.
$$

如果$A^{ik} \succ_t B^{kj}$, $\forall i, j, k$ (或$A^{ik} \prec_t B^{kj}$, $\forall i, j, k$), 那么

$$
A \ltimes B = \begin{bmatrix} C^{11} & \cdots & C^{1t} \\ \vdots & & \vdots \\ C^{r1} & \cdots & C^{rt} \end{bmatrix}, \tag{2.3.23}
$$

其中

$$
C^{ij} = \sum_{k=1}^{s} A^{ik} \ltimes B^{kj}.
$$

(3) 如果$A \succ_t B$, $B \succ_s T$, 那么$A \ltimes B \succ_{ts} T$.

(4) 设$A_1, \cdots, A_n$ 是一个有序链, 那么

$$
A_1 \ltimes A_2 \ltimes \cdots \ltimes A_n
$$

有定义.

证明 通过直接计算就可以证明(1) 和(2)、(3) 是定义的直接推论. (4) 可由(3) 得到. □

实际上, (2) 说明矩阵的分块乘法可以推广到左半张量积上, 只要将第一个矩阵的行和第二个矩阵的列按照原来两个矩阵的比例分割就行了. 而(1) 说明对(2.3.22) 中的第一个矩阵A 作任何水平分割, 或者对(2.3.22) 中的第二个矩阵B 作任何垂直分割, 然后进行矩阵的分块(左半张量)乘法, 都是被允许的.

现在我们准备证明左半张量积一些重要的基本性质.

定义 2.3.15　只要$\ltimes$ 有定义, 即矩阵有合适的维数, 那么$\ltimes$ 满足

(1) 分配律

$$\begin{cases} F \ltimes (aG \pm bH) = aF \ltimes G \pm bF \ltimes H, \\ (aF \pm bG) \ltimes H = aF \ltimes H \pm bG \ltimes H, \quad a, b \in \mathbb{R}. \end{cases}$$

(2) 结合律

$$(F \ltimes G) \ltimes H = F \ltimes (G \ltimes H).$$

证明　第 1 部分(分配律)的证明很容易, 留给读者, 我们只证明第2 部分(结合律).

首先, 我们需要说明如果F, G 和H 有使$(F \ltimes G) \ltimes H$ 有定义的合适维数, 那么此时$F \ltimes (G \ltimes H)$ 也有定义.

情形1. $F \succ G$ 且$G \succ H$. 因此可以假设F, G 和H 的维数分别是$m \times np$, $p \times qr$ 和$r \times s$.

于是$F \ltimes G$ 的维数是$m \times nqr$, $(F \ltimes G) \ltimes H$ 有定义. 另一方面$G \ltimes H$ 的维数是$p \times qs$, $F \ltimes (G \ltimes H)$ 有定义.

情形2. $F \prec G$ 且$G \prec H$. 因此, 可以假设F, G 和H 的维数分别是$m \times n$, $np \times q$ 和$rq \times s$.

$F \ltimes G$ 的维数是$mp \times q$, $(F \ltimes G) \ltimes H$ 有定义. 另一方面$G \ltimes H$ 的维数是$npr \times s$, $F \ltimes (G \ltimes H)$ 有定义.

情形3. $F \prec G$ 且$G \succ H$. 因此, 可以假设F, G 和H 的维数分别是$m \times n$, $np \times qr$ 和$r \times s$.

$F \ltimes G$ 的维数是$mp \times qr$, $(F \ltimes G) \ltimes H$ 有定义. 另一方面$G \ltimes H$ 的维数是$np \times qs$, $F \ltimes (G \ltimes H)$ 有定义.

情形4. $F \succ G$ 且$G \prec H$. 因此, 可以假设F, G 和H 的维数分别是$m \times np$, $p \times q$ 和$rq \times s$.

$F \ltimes G$ 的维数是$m \times nq$. 为了让$(F \ltimes G) \ltimes H$ 有定义, 我们需要

情形4.1. $(F \ltimes G) \succ H$, 即$n = n'r$, $F \ltimes (G \ltimes H)$ 有定义;

情形4.2. $(F \ltimes G) \prec H$, 即$r = nr'$, $F \ltimes (G \ltimes H)$ 有定义;

$G \ltimes H$ 的维数是$pr \times s$. 为了让$(F \ltimes G) \ltimes H$ 有定义, 我们需要

情形4.3. $F \succ (G \ltimes H)$, 即$n = n'r$, $(F \ltimes G) \ltimes H$ 有定义;

情形4.4. $F \prec (G \ltimes H)$, 即$r = nr'$, $(F \ltimes G) \ltimes H$ 有定义;

接下来, 我们证明结合律. 我们需要逐一对每种情形进行证明. 但是由于情形1–3 的证明是类似的, 所以我们只证明情形1, 即$F \succ G$ 且$G \succ H$ 时的情形.

给定$F_{m\times np}$, $G_{p\times qr}$, $H_{r\times s}$. 不失一般性我们假设$m = 1$, $s = 1$, 由(2.3.22) 有

$$F \ltimes G = \begin{bmatrix} F_1 & \cdots & F_p \end{bmatrix} \ltimes \begin{bmatrix} g_{11}^1 & \cdots & g_{1q}^1 & \cdots & g_{r1}^1 & \cdots & g_{rq}^1 \\ \vdots & & \vdots & & \vdots & & \vdots \\ g_{11}^p & \cdots & g_{1q}^p & \cdots & g_{r1}^p & \cdots & g_{rq}^p \end{bmatrix}$$

$$= \left[\sum_{i=1}^p F_i g_{11}^i, \quad \cdots, \quad \sum_{i=1}^p F_i g_{1q}^i, \quad \cdots, \quad \sum_{i=1}^p F_i g_{r1}^i, \quad \cdots, \quad \sum_{i=1}^p F_i g_{rq}^i \right].$$

那么

$$\begin{aligned} (F \ltimes G) \ltimes H &= (F \ltimes G) \ltimes \begin{bmatrix} h_1 \\ \vdots \\ h_r \end{bmatrix} \\ &= \left[\sum_{j=1}^r \sum_{i=1}^p F_i g_{j1}^i h_j, \quad \cdots, \quad \sum_{j=1}^r \sum_{i=1}^p F_i g_{jq}^i h_j \right]. \end{aligned} \tag{2.3.24}$$

另一方面

$$\begin{bmatrix} g_{11}^1 & \cdots & g_{1q}^1 & \cdots & g_{r1}^1 & \cdots & g_{rq}^1 \\ \vdots & & \vdots & & \vdots & & \vdots \\ g_{11}^p & \cdots & g_{1q}^p & \cdots & g_{r1}^p & \cdots & g_{rq}^p \end{bmatrix} \ltimes \begin{bmatrix} h_1 \\ \vdots \\ h_r \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^r g_{j1}^1 h_j & \cdots & \sum_{j=1}^r g_{jq}^1 h_j \\ \vdots & & \vdots \\ \sum_{j=1}^r g_{j1}^p h_j & \cdots & \sum_{j=1}^r g_{jq}^p h_j \end{bmatrix}.$$

有

$$\begin{aligned} F \ltimes (G \ltimes H) = & (F_1, \cdots, F_p) \ltimes (G \ltimes H) \\ = & \left[\sum_{j=1}^r \sum_{i=1}^p F_i g_{j1}^i h_j, \quad \cdots, \quad \sum_{j=1}^r \sum_{i=1}^p F_i g_{jq}^i h_j \right], \end{aligned}$$

由上式与(2.3.24) 有$(F \ltimes G) \ltimes H = F \ltimes (G \ltimes H)$.

情形4.1~4.4 是相似的, 因此我们也只证明情形4.1.

给定$F_{m\times npr}$, $G_{p\times q}$, $H_{rq\times s}$. 我们也假设$m=1$, $s=1$, 那么

$$F=(F_{11},\cdots,F_{1r},\cdots,F_{p1},\cdots,F_{pr}),$$

其中每一个F_{ij} 都是一个$1\times n$ 的块.

$$G=\begin{bmatrix} g_{11} & \cdots & g_{1q} \\ \vdots & & \vdots \\ g_{p1} & \cdots & g_{pq} \end{bmatrix},\quad H=(h_{11},\cdots,h_{1r},\cdots,h_{q1},\cdots,h_{qr})^{\mathrm{T}}.$$

通过计算有

$$(F\ltimes G)\ltimes H=F\ltimes(G\ltimes H)=\sum_{i=1}^{p}\sum_{j=1}^{r}\sum_{k=1}^{q}F_{ij}g_{ik}h_{kj}.$$

□

在上面的证明中, 我们将一些情形留给读者去验证.

利用定义2.3.15, 由矩阵的幂的定义有

推论 2.3.16 *设A^n 有定义, 那么*

$$A^{s+t}=A^s\ltimes A^t.$$

很自然地定义

$$A^0=1\in\mathbb{R}.$$

1 在左半张量积中是一个通用单位元, 并且可以看到这种规定与乘法的指数法则是不矛盾的.

结合指数表示和结合律, 左半张量积就变成了一个更加方便有效的工具. 下面我们给出一个说明性的例子:

例 2.3.17 回顾例2.3.8 (并参考例1.3.1). 设X 和Y 是$\mathcal{G}$ 中的两个向量, 分别将它们表示成$X=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 和$Y=(y_1,y_2,\cdots,y_n)^{\mathrm{T}}$. 将李括号记为

$$[X,Y]=\mathrm{ad}_X Y.$$

这种表示在考虑多重李括号时有很多好处, 我们可以递推地定义$\mathrm{ad}_X^{k+1}Y=[X,\mathrm{ad}_X^k Y]$, 这种形式在动态(控制) 系统分析[68]中经常用到.

利用(2.3.9), 对于有限维李代数有

$$\mathrm{ad}_X^k Y = (X^{\mathrm{T}} \ltimes M_g)^k \ltimes Y = (X^{\mathrm{T}} \ltimes M_g)^k Y. \tag{2.3.25}$$

为了展示新乘法的便利, 我们考虑$sl(2,\mathbb{R})$, 它是特殊线性群$SL(2,\mathbb{R})$ 的李代数[83], 这个李代数的一个基底是:

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

线性代数的李括号定义为$[A,B] = AB - BA$, 它也称为A 和B 的交换子. 于是我们得到交换子表如表2.3.1

表 2.3.1 $sl(2,\mathbb{R})$的交换子

	A_1	A_2	A_3
A_1	0	A_1	$-2A_2$
A_2	$-A_1$	0	A_3
A_3	$2A_2$	$-A_3$	0

根据表2.3.1 和

$$[A_i, A_j] = \alpha_{i,j}^k A_k,$$

很容易得到结构常数$\alpha_{i,j}^k$, 例如, 由上表我们有

$$[A_1, A_3] = -2A_2.$$

也就是, $\alpha_{1,3}^1 = 0$, $\alpha_{1,3}^2 = -2$, $\alpha_{1,3}^3 = 0$.

按照$\binom{i,k}{j}$ 形式(或者说对于指标(i,j,k) 按照$\binom{1,3}{2}$ 形式) 排列$\alpha_{i,j}^k$, 得到结构矩阵M_g 为

$$M_g = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 2 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix},$$

可以看出

$$\mathrm{ad}_X Y = X^{\mathrm{T}} M_g Y.$$

为了得到一个数值例子, 我们假设$X = A_1 + 2A_2 + A_3 := (1\ 2\ 1)^{\mathrm{T}}$, $Y = 3A_1 - 2A_2 + 4A_3 := (3, -2, 4)^{\mathrm{T}}$, 于是

$$\mathrm{ad}_X^k Y = (X^{\mathrm{T}} \ltimes M_g)^k Y = \begin{bmatrix} -2 & 1 & 0 \\ 2 & 0 & -2 \\ 0 & -1 & 2 \end{bmatrix}^k \begin{bmatrix} 3 \\ -2 \\ 4 \end{bmatrix}.$$

同理, 我们可以将结构常数按照$\binom{k}{i,j}$ 形式(或者$\binom{3}{1,2}$) 排列, 那么我们得到$sl(2,\mathbb{R})$ 的结构矩阵M^g 为

$$M^g = \begin{bmatrix} 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \end{bmatrix}.$$

显然在这个结构矩阵下我们有

$$\mathrm{ad}_X Y = M^g X Y.$$

设$X, Y \in \mathcal{G}$, 定义一族向量场如下:

$$\begin{cases} Z_1 = \mathrm{ad}_X Y, \\ Z_{k+1} = \mathrm{ad}_{Z_k} Y, \quad k \geq 1. \end{cases}$$

利用上面的表示, 我们有

$$Z_k = (M^g)^k X Y^k.$$

□

为了说明用结构矩阵刻画李代数的优越性, 我们考虑李代数的表现. 因为 $\mathrm{ad}_X Y = [X, Y]$ 是一个 $\mathcal{G} \times \mathcal{G} \to \mathcal{G}$ 的双线性映射, 那么对于一个固定的 X, 映射 $\mathrm{ad}_X : Y \mapsto \mathrm{ad}_x Y$ 则是一个 $\mathcal{G} \to \mathcal{G}$ 的线性映射. 这个映射称为李代数 $\mathcal{G}$ 上 X 的表现.

例 2.3.18　考虑$sl(2,\mathbb{R})$. 设

$$X = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} \in gl(2,\mathbb{R}),$$

找出它的李代数表现.

利用例2.3.17 中的结构矩阵, 容易证明

$$\mathrm{ad}_X = X^{\mathrm{T}} M_g \quad 或 \quad \mathrm{ad}_X = M^g X.$$

这里X 应当表示为一个向量形式, 即$X=(2,2,3)^{\mathrm{T}}$. 直接计算可知

$$\mathrm{ad}_X = X^{\mathrm{T}} M_g = M^g X = \begin{bmatrix} -2 & 2 & 0 \\ 6 & 0 & -4 \\ 0 & -3 & 2 \end{bmatrix}.$$

□

注 我们已经提到矩阵的左半张量积是普通矩阵乘法的推广. 即, 设$A \in M_{m\times n}$, $B \in M_{p\times q}$. 如果$n=p$, 那么

$$A \ltimes B = AB.$$

因此, 在上述例子中, 并且在以后的各章中, 除非为了强调左半张量积, 否则我们将会省略乘法符号$\ltimes$. 所有省略符号的矩阵乘法都看作是左半张量积, 普通的矩阵乘法只是它的一种特殊情况.

最后,考虑列(行)向量族的左半张量积. 设$X_i \in \mathbb{R}^{n_i}$, $i=1,\cdots,k$ 为一组列(行)向量. 那么, 根据定义容易看出它们的左半张量积$X_1\cdots X_k$ 显然是定义好的. 一个在实际应用中十分重要的问题是, 我们能否改变因子顺序? 实际上, 一组列向量的积构成一个多指标列向量, 因为, 改变因子顺序相当于指标排序变换. 因此, 可利用定理1.5.8 证明如下命题.

命题 2.3.19 (1) 设$X_i \in \mathbb{R}^{n_i}$, $i=1,\cdots,k$ 为一组列向量, 则

$$\begin{aligned} &\left[I_{n_1+\cdots+n_{t-1}} \otimes W_{[n_t,n_{t+1}]} \otimes I_{n_{t+2}+\cdots+n_k}\right] X_1X_2\cdots X_k \\ &= X_1\cdots X_{t-1}X_{t+1}X_tX_{t+2}\cdots X_k. \end{aligned} \tag{2.3.26}$$

(2) 设$\omega_i \in \mathbb{R}^{n_i}$, $i=1,\cdots,k$ 为一组行向量, 则

$$\begin{aligned} &\omega_1\omega_2\cdots\omega_k\left[I_{n_{t+2}+\cdots+n_k} \otimes W_{[n_t,n_{t+1}]} \otimes I_{n_1+\cdots+n_{t-1}}\right] \\ &= \omega_1\cdots\omega_{t-1}\omega_{t+1}\omega_t\omega_{t+2}\cdots\omega_k. \end{aligned} \tag{2.3.27}$$

证明 设 X_j 的分量用 i_j 标注, $j=1,\cdots,k$. 那么, 乘积正好是用索引 $Id(i_1,\cdots,i_k;n_1,\cdots,n_k)$ 标注. 因此, 可直接应用定理1.5.8 的第一部分.

设 ω_j 的分量用 i_j 标注, $j=1,\cdots,k$. 则乘积 $\omega_1\cdots\omega_k$ 是用 $Id(i_k,i_{k-1},\cdots,i_1;n_k,n_{k-1},\cdots,n_1)$ 标注. 因此不能直接用定理1.5.8 的第二部分. 令

$$I_j=i_{k+1-j},\quad N_j=n_{k+1-j},\quad j=1,\cdots,k,$$

$$s:=k-t,\quad s+1=k+1-t,$$

则ω_t 与ω_{t+1} 的交换变为指标$s+1$ 与s 的交换. 对新指标应用定理1.5.8 的第二部分, 转换矩阵应为

$$I_{N_1+\cdots+N_{s-1}}\otimes W_{[N_{s+1},N_s]}\otimes I_{N_{s+2}+\cdots+N_k},$$

根据定义, 它就是

$$I_{n_k+\cdots+n_{t+2}}\otimes W_{[n_t,n_{t+1}]}\otimes I_{n_{t-1}+\cdots+n_1}.$$

因此可得(2.3.27). □

2.4 双线性映射

当半张量积用于两个向量时, 它变成一个双线性映射, 本节对这个映射作一些初步的讨论.

定义 2.4.1 设E, F, G 为3 个向量空间, 映射$\phi:E\times F\to G$ 称为一个双线性映射, 如果

$$\phi(aX_1+bX_2,Y)=a\phi(X_1,Y)+b\phi(X_2,Y),$$

$$\phi(X,cY_1+dY_2)=c\phi(X,Y_1)+d\phi(X,Y_2), \tag{2.4.1}$$

$$a,b,c,d\in\mathbb{R},\quad X,X_1,X_2\in E,\quad Y,Y_1,Y_2\in F.$$

设$\{e_1,\cdots,e_m\}$ 和$\{f_1,\cdots,f_n\}$ 分别为E 和F 的基底. 定义

$$t_{i,j}=\phi(e_i,f_j),\quad 1\le i\le m;1\le j\le n;$$

$$T=\mathrm{Span}\{t_{i,j}\}\subset G.$$

则显见T 是包含象集$Im(\phi)$ 的最小子空间, 记之为$T=IM(\phi)$.

设

$$\{t_{i,j}\mid 1\le i\le m;1\le j\le n\}$$

线性无关,则它们成为T 的一组基. 我们用索引$Id(i,j;m,n)$ 排序,那么$\phi : E\times F \to T$ 就可以用坐标形式表示为

$$\phi(X,Y) = X \ltimes Y.$$

首先, 我们注意到, 在一般情况下$\mathrm{Im}(\phi) \neq \mathrm{IM}(\phi)$, 或者说$\mathrm{Im}(\phi)$ 不是一个子空间. 例如, 取$E = F = \mathbb{R}^2$, 其基底为标准基底$\{\delta_1^2, \delta_2^2\}$. 取$G = \mathbb{R}^4$, 其基底为标准基底$\{\delta_1^4, \delta_2^4, \delta_3^4, \delta_4^4\}$. 取半张量积为其双线性映射, 容易看出, 它的象为

$$\mathrm{Im}(\phi) = \left\{ z \in \mathbb{R}^4 \;\middle|\; z_1 z_4 = z_2 z_3 \right\}.$$

从几何角度看, 它不是一个3 维流形, 它可以由两个部分构成: 当$z_1 \neq 0$ 时, z_4 可由上式唯一确定. 因此, 它可看作在3 维空间z_1-z_2-z_3 中将$z_1 = 0$ 平面除去, 即$\mathbb{R}^3\backslash\{z_1 = 0\}$. 而当$z_1 = 0$ 时, 它可看作在3 维空间z_2-z_3-z_4 中$\{z_2 = 0\}$ 及$\{z_3 = 0\}$ 平面的交.

一般地说, 当我们把半张量积看作$\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^{mn}$ 的双线性映射时, 它的象用双指标记作$\{z_{i,j}\}$, 满足

$$z_{i,p}z_{j,q} = z_{i,q}z_{j,p}, \quad 1 \leq i \neq j \leq m;\ 1 \leq p \neq q \leq n.$$

约束方程的个数为

$$\binom{m}{2}\binom{n}{2} = \frac{m(m-1)n(n-1)}{4}.$$

定义 2.4.2 称一个双线性映射$\otimes : E \times F \to T$ 为泛映射, 如果对任一双线性映射$\phi : E \times F \to H$, $T = IM(\otimes)$, 均存在唯一的$f : T \to H$, 使得图2.4.1 可交换, 即

$$\phi = f \circ \otimes.$$

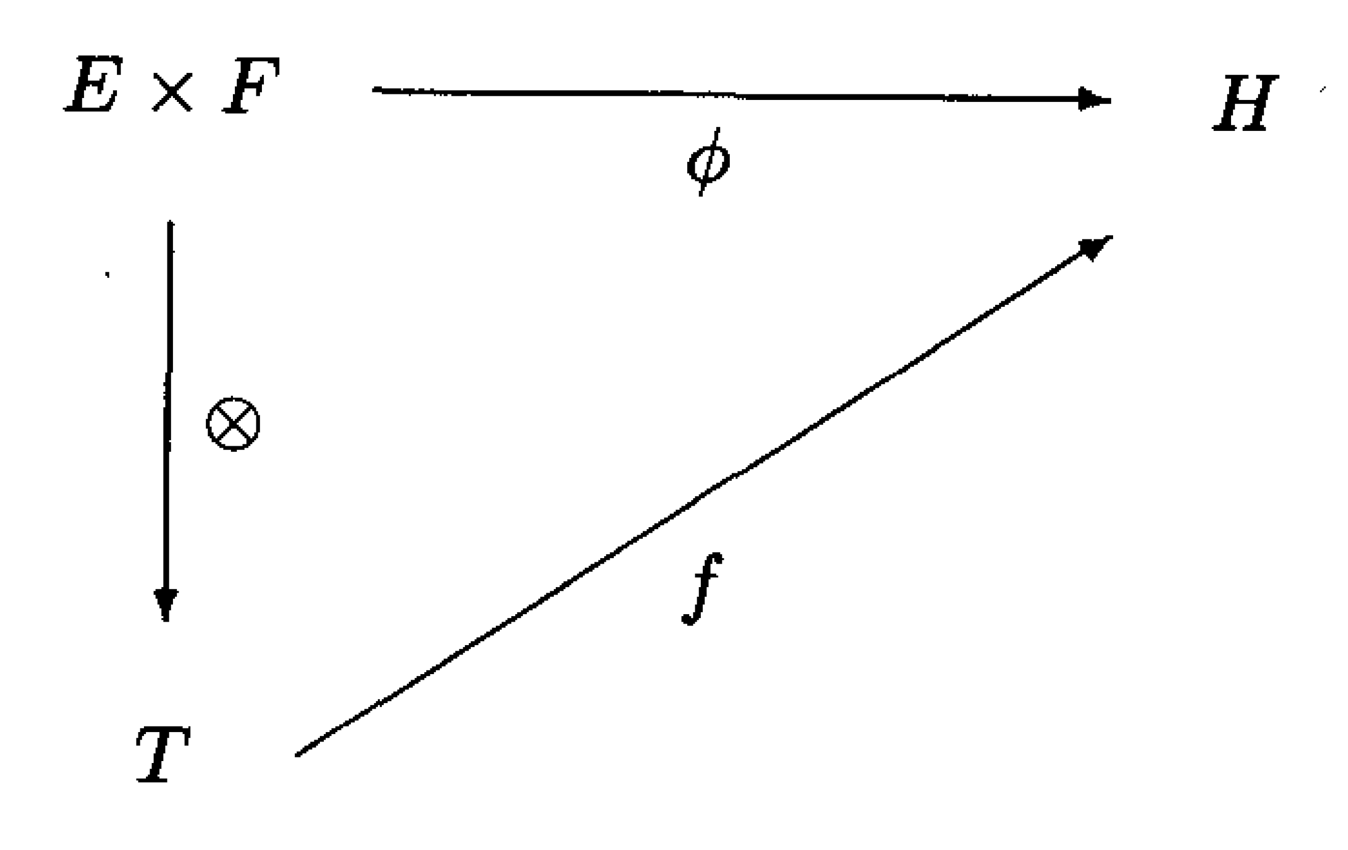

图 2.4.1 双线性泛映射

命题 2.4.3　半张量积$\ltimes : E \times F \to T$ 是一个双线性泛映射.

证明　设$\phi : E \times F \to H$ 为一双线性映射, 那么

$$\phi(X,Y) = M_\phi XY, \quad X \in E,\ Y \in F.$$

这里, M_ϕ 是ϕ 的结构矩阵. 定义$f : T \to H$ 为线性映射$Z \mapsto M_\phi Z$, 则得$\phi = f \circ \ltimes$. f 的唯一性来自双线性映射结构矩阵的唯一性. □

下面证明半张量积作为双线性映射的一个有用性质.

命题 2.4.4　设$X_1, \cdots, X_k \in E$, $Y_1, \cdots, Y_k \in F$, 并且

$$\sum_{i=1}^{k} X_i \ltimes Y_i = 0.$$

(1) 如果$Y_1, \cdots, Y_k$ 线性无关, 则$X_1 = \cdots = X_k = 0$.
(2) 如果$X_1, \cdots, X_k$ 线性无关, 则$Y_1 = \cdots = Y_k = 0$.

证明　(1)

$$\sum_{i=1}^{k} X_i \ltimes Y_i = \begin{bmatrix} \sum\limits_{i=1}^{k} x_1^i Y_i \\ \vdots \\ \sum\limits_{i=1}^{k} x_m^i Y_i \end{bmatrix} = 0.$$

结论显见.

(2)

$$\sum_{i=1}^{k} Y_i \ltimes X_i = W_{[m,n]} \sum_{i=1}^{k} X_i \ltimes Y_i = 0.$$

利用上面的结论即得. □

2.5　注释与参考

矩阵的张量积, 或称Kronecker 积, 已经有十分广泛的应用, 它的性质和应用可在许多标准参考书上找到, 例如文献[65]. 矩阵的Hadamard 积可参见文献[124].

关于立方阵的基本概念和应用可见文献[107], [5], [97], [108]. 关于立方阵与半张量积的关系的更详细的讨论可见文献[125].

关于双线性或多线性映射及其性质, 建议参考文献[54]. 实际上,半张量积可以在多线性代数中找到许多应用. 笔者相信这方面有许多研究课题可以做.

习 题 二

2.1 证明: 如果$A \in M_{m\times n}$, $B \in M_{n\times s}$, 那么左半张量积就是普通的矩阵乘法, 即

$$A \ltimes B = AB.$$

2.2 证明命题2.1.4.

2.3 证明命题2.2.7.

2.4 证明命题2.3.10.

2.5 设$x = (x_1, x_2, x_3) \in \mathbb{R}^3$.

(1) 请将齐次多项式

$$F(x) = 2x_1^3 - 2x_1x_3^2 + x_2^2x^3,$$

表示成

$$F(x) = Cx^3.$$

(2) 请将多项式

$$Q(x) = x_1x_2 + x_1x_3 + x_2x_3 + x_1^2x_2 + x_2^2x_3 + x_3^2x_1$$

表示成

$$Q(x) = Ax^2 + Bx^3.$$

2.6 设

$$A = \begin{bmatrix} 1 & 2 & -1 & 0 \\ 0 & 1 & 2 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 1 & 1 & 0 & -1 \end{bmatrix}.$$

计算(1) AB; (2) A^2; (3) $A^{\mathrm{T}(2\times 2)}B$; (4) AB^2.

2.7 考虑$\mathbb{R}^3$ 中的叉乘“×”. 取$\mathbb{R}^3$ 中的一组基底$\{I, J, K\}$, 叉乘定义如下: 令$X = x_1I + x_2J + x_3K := (x_1, x_2, x_3)^{\mathrm{T}}$, $Y = y_1I + y_2J + y_3K := (y_1, y_2, y_3)^{\mathrm{T}}$, $Z = z_1I + z_2J + z_3K := (z_1, z_2, z_3)^{\mathrm{T}}$, 定义$Z = X \times Y$ 为

$$Z = \det \begin{bmatrix} I & J & K \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}.$$

(1) 找出结构常数$\alpha_{i,j}^k$;

(2) 按照$\binom{ik}{j}$ 形式排列$\alpha_{i,j}^k$ 得到结构矩阵$M^{\times}$, 并且证明

$$X \times Y = X^{\mathrm{T}}M^{\times}Y.$$

(3) 按照$\binom{k}{ij}$ 形式排列$\alpha_{i,j}^k$ 得到结构矩阵$M^{\times}$, 并且证明

$$X \times Y = M^{\times} XY.$$

2.8 设$X=(1,2,-1)^{\mathrm{T}}$, $Y=(2,2,0)^{\mathrm{T}}$. 利用习题2.7 的结果计算
(1) $X\times Y$;
(2) $\underbrace{X\times X\times\cdots\times X}_{100}\times Y$.

2.9 设$X\in\mathbb{R}^3$. 它的表现$\mathrm{ad}_X:\mathbb{R}^3\to\mathbb{R}^3$ 定义为$Y\mapsto X\times Y$. 请给出它的表现ad_X 的矩阵形式.

2.10 (1) 设列向量$X\in\mathbb{R}^m$, $Y\in\mathbb{R}^n$, 证明

$$W_{[m,n]}XY = YX;$$

(2) 设行向量$\omega\in\mathbb{R}^m$, $\mu\in\mathbb{R}^n$, 证明

$$\omega\mu W_{[m,n]} = \mu\omega.$$

2.11 (1) 设列向量$X_i\in\mathbb{R}^{n_i}$, $i=1,\cdots,k$, 证明

$$\left(I_{n_1+\cdots+n_{t-1}}\otimes W_{[n_t,n_{t+1}]}\right)X_1\cdots X_tX_{t+1}\cdots X_k = X_1\cdots X_{t+1}X_t\cdots X_k;$$

(2) 设行向量$\omega_i\in\mathbb{R}^{n_i}$, $i=1,\cdots,k$, 证明

$$\omega_1\cdots\omega_t\omega_{t+1}\cdots\omega_k\left(I_{n_1+\cdots+n_{t-1}}\otimes W_{[n_t,n_{t+1}]}\right) = \omega_1\cdots\omega_{t+1}\omega_t\cdots\omega_k.$$

比较它们与定理1.5.8 的异同点.

2.12 令$\sigma\in\mathbf{S}_k$ 为一k 阶置换, $n=n_1\times\cdots\times n_k$, $n_i\geq 1$. 定义$n\times n$ 矩阵W_σ 如下: 用k 个指标按索引$Id(i_1,\cdots,i_k;n_1,\cdots,n_k)$ 标注其行, 用k 个指标按索引$Id(j_{\sigma_1},\cdots,j_{\sigma_k};n_{\sigma_1},\cdots,n_{\sigma_k})$ 标注其列, 然后定义其第$i_1\cdots i_k$ 列第$j_1\cdots j_k$ 行的元素为

$$\omega_{j_1,\cdots,j_k}^{i_1,\cdots,i_k} = \begin{cases}1, & i_1=j_1,\cdots,i_k=j_k;\\ 0, & \text{其他}.\end{cases}$$

称W_σ 为σ 置换矩阵.

(1) 设列向量$X_i\in\mathbb{R}^{n_i}$, $i=1,\cdots,k$, 证明

$$W_\sigma X_1\cdots X_k = X_{\sigma_1}\cdots X_{\sigma_k};$$

(2) 设行向量$\omega_i\in\mathbb{R}^{n_i}$, $i=1,\cdots,k$, 证明

$$\omega_1\cdots\omega_kW_\sigma = \omega_{\sigma_1}\cdots\omega_{\sigma_k}.$$

2.13 设$\sigma = \sigma_k\sigma_{k-1}\cdots\sigma_1$, 那么

$$W_\sigma = W_{\sigma_k}W_{\sigma_{k-1}}\cdots W_1.$$

2.14 设$n_1 = n_2 = n_3 = 2$, $\sigma = (1,2,3) \in \mathbf{S}_3$.

(1) 构造一个置换矩阵W_σ;

(2) 令$\sigma_1 = (12)$, $\sigma_2 = (13)$, 则$\sigma = \sigma_2\sigma_1$. 用换位矩阵构造$W_{\sigma_1}$ 和W_{σ_2}. 验证

$$W_\sigma = W_{\sigma_2}W_{\sigma_1}.$$

(3)证明任何一个置换矩阵均可由换位矩阵生成.

2.15 设$X_i \in \mathbb{R}^u$, $Y_{ij} \in \mathbb{R}^v$, $Z_{ijk} \in \mathbb{R}^s$, $W_{ijk} \in \mathbb{R}^t$. 设$\{X_i \,|\, 1 \le i \le \alpha\}$ 线性无关, $\{Y_{ij} \,|\, 1 \le i \le \alpha; 1 \le j \le \beta\}$ 线性无关, $\{Z_{ijk} \,|\, 1 \le i \le \alpha; 1 \le j \le \beta; 1 \le k \le \gamma\}$ 线性无关, 且

$$\sum_{i=1}^{\alpha}\sum_{i=1}^{\alpha}\sum_{j=1}^{\beta}\sum_{k=1}^{\gamma} X_iY_{ij}Z_{ijk}W_{ijk} = 0.$$

则

$$W_{ijk} = 0, \quad 1 \le i \le \alpha; 1 \le j \le \beta; 1 \le k \le \gamma.$$

试证之.

第 3 章　左半张量积与矩阵映射

本章主要研究矩阵左半张量积的性质. 首先证明, 几乎所有普通矩阵乘法的重要性质都可以推广到左半张量积. 接着, 利用左半张量积研究矩阵的线性和多项式映射. 再根据矩阵映射的性质得出左半张量积的一些新性质. 最后, 讨论了普通矩阵乘法、张量积和左半张量积之间的转化.

3.1　基本性质

本节讨论左半张量积的一些基本性质. 不难看出, 当普通矩阵乘法推广到左半张量积时, 几乎所有的乘法性质都保留下来了. 左半张量积的生命力正在于此.

命题 3.1.1　设A 和B 是两个具有合适维数的矩阵, 那么

$$(A \ltimes B)^{\mathrm{T}} = B^{\mathrm{T}} \ltimes A^{\mathrm{T}}. \tag{3.1.1}$$

证明　通过简单计算可知, 对于具有合适维数的行向量X 和列向量Y, 有

$$\langle X, Y\rangle_L = \left(\langle Y^{\mathrm{T}}, X^{\mathrm{T}}\rangle_L\right)^{\mathrm{T}}. \tag{3.1.2}$$

考虑$A \ltimes B$. 记A 的第i 行为A^i, B 的第j 列为B_j, 那么显然$A \ltimes B$ 的第(i,j)块就是

$$\left\langle A^i, B_j\right\rangle_L,$$

此时$B^{\mathrm{T}} \ltimes A^{\mathrm{T}}$的第$(j,i)$块是

$$\left\langle B_j^{\mathrm{T}}, (A^i)^{\mathrm{T}}\right\rangle_L.$$

由(3.1.2), 可以看出$A \ltimes B$ 的第(i,j) 块的转置就是$B^{\mathrm{T}} \ltimes A^{\mathrm{T}}$ 的第(j,i) 块, 于是命题得证. □

下面的命题说明两个矩阵的左半张量积可以很容易地用它们的普通积加上Kronecker 积来实现:

命题 3.1.2　(1) 如果$A \in M_{m\times np}$, $B \in M_{p\times q}$, 那么

$$A \ltimes B = A(B \otimes I_n). \tag{3.1.3}$$

(2) 如果$A \in M_{m\times n}$, $B \in M_{np\times q}$, 那么

$$A \ltimes B = (A \otimes I_p)B. \tag{3.1.4}$$

证明 根据命题2.3.14, 不失一般性, 对于矩阵A 和B 我们可以假设$m = q = 1$, 那么可以通过直接计算验证等式成立. □

命题3.1.2 是很基本的结论, 半张量积的很多性质都可以由它得到. 下面的命题可以认为是命题3.1.2 的直接推论.

命题 3.1.3 给定两个具有合适维数的方阵A 和B, 并且$A \ltimes B$ 有定义, 那么

(1) $A \ltimes B$ 和$B \ltimes A$ 有相同的特征函数.

(2) $\mathrm{tr}(A \ltimes B) = \mathrm{tr}(B \ltimes A)$.

(3) 如果A 或B 可逆, 那么$A \ltimes B \sim B \ltimes A$, 这里"$\sim$"表示矩阵相似.

(4) 如果A 和B 都是上三角阵(下三角阵, 对角阵, 正交阵), 那么$A \ltimes B$ 也同样是三角阵(下三角阵, 对角阵, 正交阵).

(5) 如果A 和B 都可逆, 那么$A \ltimes B$ 也可逆, 并且

$$(A \ltimes B)^{-1} = B^{-1} \ltimes A^{-1}. \tag{3.1.5}$$

(6) 如果$A \prec_t B$, 那么

$$\det(A \ltimes B) = [\det(A)]^t \det(B). \tag{3.1.6}$$

如果$A \succ_t B$, 那么

$$\det(A \ltimes B) = \det(A)[\det(B)]^t. \tag{3.1.7}$$

证明 利用(3.1.3) 和(3.1.4) 将左半张量积转化为矩阵普通乘法和Kronecker积的形式, 很容易就得到上面的性质. 我们证明(5) 作为示例. 设$A \prec B$, 那么

$$(A \ltimes B)^{-1} = (A(B \otimes I))^{-1} = (B \otimes I)^{-1}A^{-1} = (B^{-1} \otimes I)A^{-1} = B^{-1} \ltimes A^{-1}.$$

□

下面的命题表明换位矩阵也可以交换块结构数组的各块的位置:

命题 3.1.4 (1) 设

$$A = (A_{11}, \cdots, A_{1n}, \cdots, A_{m1}, \cdots, A_{mn})$$

是每个分块都有相同维数的矩阵, 它是由指标$\{i, j\}$ 按照索引$Id(i, j; m, n)$ 排列的, 那么

$$AW_{[n,m]} = (A_{11}, \cdots, A_{m1}, \cdots, A_{1n}, \cdots, A_{mn})$$

是按照索引$Id(j, i; n, m)$ 排列的.

(2) 设

$$B = \mathrm{col}(B_{11}, \cdots, B_{1n}, \cdots, B_{m1}, \cdots, B_{mn})$$

是由具有相同维数的分块排成一列的矩阵, 由$\{i,j\}$ 按照索引$Id(i,j;m,n)$ 排列, 那么

$$W_{[m,n]}B = \mathrm{col}(B_{11}, \cdots, B_{m1}, \cdots, B_{1n}, \cdots, M_{mn})$$

是按照索引$Id(j,i;n,m)$ 排列的.

证明 如果A_{ij} 是列向量或者B_{ij} 是行向量, 由命题1.5.7 可直接得到结果(见习题1.13).

利用命题3.1.2, 可以看出根据左半张量积, 换位矩阵也可以实现分块的重新排列. □

一个矩阵和单位阵I_n 的左半张量积有一些特殊的性质. 粗略地说, 当I 的大小小于或等于矩阵M 的大小(这里, 大小指的是行数或列数, 分别对应于I 左乘或右乘M) 时, 它就是一个单位阵. 当I 的大小大于M 的大小时, 它将会扩大M.

命题 3.1.5 (1) 设M 是一个$m \times pn$ 矩阵, 那么

$$M \ltimes I_n = M. \tag{3.1.8}$$

(2) 设M 是一个$m \times n$ 矩阵, 那么

$$M \ltimes I_{pn} = M \otimes I_p. \tag{3.1.9}$$

(3) 设M 是一个$pm \times n$ 矩阵, 那么

$$I_p \ltimes M = M. \tag{3.1.10}$$

(4) 设M 是一个$m \times n$ 矩阵, 那么

$$I_{pm} \ltimes M = M \otimes I_p. \tag{3.1.11}$$

证明 所有的等式都可以利用命题3.1.2 直接推导出来. (我们将具体的验证留给读者.) □

下面的命题说明左半张量积可以用来将一些有关矩阵的线性映射表示成它们的展开式的线性映射. 在下一节里我们将会讨论另一种表示.

命题 3.1.6 设$A \in M_{m\times n}$, $X \in M_{n\times q}$, $Y \in M_{p\times m}$, 那么

$$V_r(AX) = A \ltimes V_r(X), \tag{3.1.12}$$

$$V_c(YA) = A^{\mathrm{T}} \ltimes V_c(Y). \tag{3.1.13}$$

证明 对于(3.1.12), 令$C = AX$, 并且记A 的第i 行为A^i. 根据命题2.3.14, 等式右边的第i 块就是

$$
\begin{aligned}
A^i \ltimes V_r(X) &= \left\langle A^i, (x_{11}, \cdots, x_{1s}, \cdots, x_{n1}, \cdots, x_{ns})^{\mathrm{T}} \right\rangle_L \\
&= \begin{bmatrix} \sum_{k=1}^{n} a_{ik} x_{k1} \\ \vdots \\ \sum_{k=1}^{n} a_{ik} x_{ks} \end{bmatrix} = (C^i)^{\mathrm{T}},
\end{aligned}
$$

于是(3.1.12) 成立. 再来证明(3.1.13), 对(3.2.11) 应用(1.5.4), 就有

$$
V_c(YA) = V_r(A^{\mathrm{T}} Y^{\mathrm{T}}) = A^{\mathrm{T}} \ltimes V_r(Y^{\mathrm{T}}) = A^{\mathrm{T}} \ltimes V_c(Y).
$$

□

注意到(3.1.12) 形如一个向量空间(例如$\mathbb{R}^n$ 等) 上的线性映射. 实际上, 当X 是一个向量时, (3.1.12) 就变成一个线性映射的标准形式. 这也从另一方面表明左半张量积是普通矩阵乘法的推广.

利用(3.1.12) 和(3.1.13) 可以得到一个矩阵多项式的矩阵表示, 下面就是一个直接结果. 我们将其详细证明留给读者.

推论 3.1.7 设X 是一个方阵, $p(x)$ 是一个多项式. 那么$p(x)$ 可以表示成$p(x) = q(x)x + p_0$ 的形式, 并且

$$
V_r(p(X)) = q(X)V_r(X) + p_0 V_r(I). \tag{3.1.14}
$$

3.2 矩阵的映射

在第 1 章中我们已经讨论了矩阵的行展开和列展开, 有时将矩阵表示成向量形式会给我们带来很多方便. 在本节, 我们将考虑矩阵表示成向量时, 一个矩阵函数(特别是线性函数) 的表示. 我们称之为矩阵映射的展开表示.

从向量空间V 到向量空间W 的线性映射: $\phi : V \to W$ 是满足线性条件的映射, 即

$$
\phi(\alpha v_1 + \beta v_2) = \alpha\phi(v_1) + \beta\phi(v_2), \quad v_1, v_2 \in V, \ \alpha, \beta \in \mathbb{R}.
$$

为方便, 我们记$L(V, W)$ 为由向量空间V 到向量空间W 的线性映射全体. 作为特例, $p \times q$ 矩阵集合到$m \times n$ 矩阵集合的线性映射的全体记作$L(M_{p\times q}, M_{m\times n})$.

我们从几个例子开始:

例 3.2.1　(1) (Lyapunov 映射) 给出一个方阵$A \in M_n$. 考虑如下映射$L_A : M_n \to M_n$, 定义为

$$L_A(X) = AX + XA^{\mathrm{T}}. \tag{3.2.1}$$

说一个矩阵是Hurwitz 阵, 如果它的所有特征值都具有负实部. 大家都知道[109]: A 是一个Hurwitz 阵, 当且仅当, 对于任意一个负定矩阵$Q < 0$, $L_A(X) = Q$ 有一个正定解. 作为向量空间M_n 上的线性映射, L_A 有一个矩阵表示[85]

$$M_{L_A}^c = A \otimes I + I \otimes A. \tag{3.2.2}$$

这种矩阵表示的确切含义就是

$$V_c(L_A(X)) = M_{L_A}^c V_c(X). \tag{3.2.3}$$

(2) (辛映射) 矩阵$X \in \mathrm{sp}(2n, \mathbb{R})$, 当且仅当, $JX + X^{\mathrm{T}}J = 0$, 其中

$$J = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$

一般地, 我们考虑广义辛映射[15]

$$L_N(X) = NX + X^{\mathrm{T}}N.$$

可以证明$L_N(X) = 0$ 的解构成一个李代数, 并且是李群

$$G_N = \{Z \in GL(n, R) \mid Z^{\mathrm{T}}NZ = N\}$$

的李代数, 而且, 对于任意的N, $L_N(X) = 0$ 存在一个非零解X. 即0 为L_N 的一个特征值.
□

在(3.2.3) 中, 矩阵表示成列展开形式, 我们使用上标c 来表示. 同样当矩阵表示成行展开时, 我们也有一个矩阵表示, 即

$$V_r(L_A(X)) = M_{L_A}^r V_r(X). \tag{3.2.4}$$

下面的命题表明, 上述两种矩阵表示可以很容易地互相转换, 即其中一种矩阵表示可由另一种得到.

命题 3.2.2　设$\rho \in L(M_{p\times q}, M_{m\times n})$, 那么

$$\begin{cases} M_\rho^r = W_{[n,m]} M_\rho^c W_{[p,q]}, \\ M_\rho^c = W_{[m,n]} M_\rho^r W_{[q,p]}. \end{cases} \tag{3.2.5}$$

特别, 如果$\rho \in L(M_n)$, 那么(3.2.5) 就变成

$$\begin{cases} M_\rho^r = W_{[n]} M_\rho^c W_{[n]}, \\ M_\rho^c = W_{[n]} M_\rho^r W_{[n]}. \end{cases} \tag{3.2.6}$$

证明 考虑(3.2.5) 中的第一式: 根据命题1.5.4, 我们可以通过三个步骤得到M_ρ^r. 首先, 用$W_{[p,q]}$ 将$X_{p\times q}$ 的行展开转化为列展开; 然后作映射M_ρ^c. 最后, 用$W_{[n,m]}$ 将$\rho(X)_{m\times n}$ 的列展开转化为行展开.

第二式可由第一式利用(1.5.5)得到. □

根据(3.2.5), M^r 和M^c 中的任一个都可以轻易地由另一个得到. 在这里我们约定, 如果没有特别说明, 当说展开时指的都是M^c.

设$Z \in M_{n\times p}$, 我们考虑一般的线性映射$\rho : M_{n\times p} \to M_{m\times q}$, 定义如下:

$$Z \mapsto AZB + CZ^{\mathrm{T}}D, \tag{3.2.7}$$

其中$A \in M_{m\times n}$, $B \in M_{p\times q}$, $C \in M_{m\times p}$, $D \in M_{n\times q}$.

对于这种一般形式我们有

命题 3.2.3 (3.2.7) 的矩阵表示是

$$M^c = (B^{\mathrm{T}} \otimes A) + (D^{\mathrm{T}} \otimes C)W_{[p,n]}. \tag{3.2.8}$$

证明 首先列出四种基本线性映射的矩阵表示(相应矩阵的维数同上),见表3.2.1.

表 3.2.1 列展开表示

ρ:	M_ρ^c
$Z \mapsto AZ$	$I_p \otimes A$
$Z \mapsto ZB$	$B^{\mathrm{T}} \otimes I_n$
$Z \mapsto CZ^{\mathrm{T}}$	$(I_n \otimes C)W_{[p,n]}$
$Z \mapsto Z^{\mathrm{T}}D$	$(D^{\mathrm{T}} \otimes I_p)W_{[p,n]}$

前两种映射是大家比较熟悉的, 它们都可以通过直接计算验证. 对于第三种映射, 有

$$\begin{aligned} V_c(CZ^{\mathrm{T}}) &= (I_n \otimes C)V_c(Z^{\mathrm{T}}) = (I_n \otimes C)V_r(Z) \\ &= (I_n \otimes C)W_{[p,n]}W_{[n,p]}V_r(Z) = (I_n \otimes C)W_{[p,n]}V_c(Z). \end{aligned}$$

而对于最后一个, 可以证明

$$\begin{aligned} V_c(Z^{\mathrm{T}}D) &= (D^{\mathrm{T}} \otimes I_p)V_c(Z^{\mathrm{T}}) = (D^{\mathrm{T}} \otimes I_p)V_r(Z) \\ &= (D^{\mathrm{T}} \otimes I_p)W_{[p,n]}W_{[n,p]}V_r(Z) = (D^{\mathrm{T}} \otimes I_p)W_{[p,n]}V_c(Z). \end{aligned}$$

将(3.2.7) 看作是两个复合线性映射的线性组合, 由上面这些映射的矩阵表示就可以得到(3.2.7) 每一项的矩阵表示. 另外注意到, 对于复合映射, 它的维数需要根据乘积中间的矩阵大小来调整, 于是我们有

$$M^c = (B^{\mathrm{T}} \otimes I_m)(I_p \otimes A) + (D^{\mathrm{T}} \otimes I_m)(I_n \otimes C)W_{[p,n]}.$$

立即就得到了(3.2.8). □

下面我们给出一个数值例子说明上面的公式:

例 3.2.4　设$A, C \in M_{3\times 2}$, $B, D \in M_{2\times 4}$, 且

$$A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & -1 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 0 & 2 & 1 \end{bmatrix}.$$

(1) 设$\rho : M_2 \to M_{3\times 2}$, 使得$Z \mapsto AZ$, 那么

$$M_\rho = I_2 \otimes A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

因此

$$V_c(AZ) = M_\rho V_c(Z) = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \\ z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} z_{11} - z_{21} \\ 2z_{11} + z_{21} \\ z_{21} \\ z_{12} - z_{22} \\ 2z_{12} + z_{22} \\ z_{22} \end{bmatrix}. \tag{3.2.9}$$

直接计算有

$$AZ = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} = \begin{bmatrix} z_{11} - z_{21} & z_{12} - z_{22} \\ 2z_{11} + z_{21} & 2z_{12} + z_{22} \\ z_{21} & z_{22} \end{bmatrix}.$$

这就说明(3.2.9)是正确的.

(2) 设$\rho : M_2 \to M_{2\times 4}$, 使得$Z \mapsto ZB$, 那么

$$M_\rho = B^{\mathrm{T}} \otimes I_2 = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

因此

$$\begin{aligned} V_c(ZB) &= M_\rho V_c(Z) \\ &= \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \\ z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} z_{11} + 2z_{12} \\ z_{21} + 2z_{22} \\ -z_{11} + z_{12} \\ -z_{21} + z_{22} \\ 2z_{12} + z_{22} \\ z_{12} \\ z_{22} \\ z_{11} \\ z_{21} \end{bmatrix}. \end{aligned} \tag{3.2.10}$$

直接计算得到

$$\begin{aligned} ZB &= \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} z_{11} + 2z_{12} & -z_{11} + z_{12} & z_{12} & z_{11} \\ z_{21} + 2z_{22} & -z_{21} + z_{22} & z_{22} & z_{21} \end{bmatrix}. \end{aligned}$$

这验证了(3.2.10) 的正确性.

(3) 设$\rho : M_2 \to M_{3\times 2}$, 使得$Z \mapsto CZ^{\mathrm{T}}$. 注意到

$$
\begin{array}{c}
\\
W_{[2,2]} =
\end{array}
\begin{array}{c}
\begin{array}{cccc}(11) & (12) & (21) & (22)\end{array}\\
\left[\begin{array}{cccc}
1 & 0 & 0 & 0\\
0 & 0 & 1 & 0\\
0 & 1 & 0 & 0\\
0 & 0 & 0 & 1
\end{array}\right]
\end{array}
\begin{array}{c}
\\
(11)\\ (21)\\ (12)\\ (22)
\end{array},
$$

于是

$$
M_\rho = (I_2 \otimes C)W_{[2,2]} = \begin{bmatrix}
1 & 0 & 2 & 0\\
0 & 0 & 1 & 0\\
-1 & 0 & 1 & 0\\
0 & 1 & 0 & 2\\
0 & 0 & 0 & 1\\
0 & -1 & 0 & 1
\end{bmatrix}.
$$

因此

$$
\begin{aligned}
V_c(CZ^{\mathrm{T}}) &= M_\rho V_c(Z)\\
&= \begin{bmatrix}
1 & 0 & 2 & 0\\
0 & 0 & 1 & 0\\
-1 & 0 & 1 & 0\\
0 & 1 & 0 & 2\\
0 & 0 & 0 & 1\\
0 & -1 & 0 & 1
\end{bmatrix}
\begin{bmatrix} z_{11}\\ z_{21}\\ z_{12}\\ z_{22}\end{bmatrix}
= \begin{bmatrix}
z_{11}+2z_{12}\\
z_{12}\\
-z_{11}+z_{12}\\
z_{21}+2z_{22}\\
z_{22}\\
-z_{21}+z_{22}
\end{bmatrix}.
\end{aligned} \tag{3.2.11}
$$

直接计算得

$$
CZ^{\mathrm{T}} = \begin{bmatrix} 1 & 2\\ 0 & 1\\ -1 & 1\end{bmatrix}
\begin{bmatrix} z_{11} & z_{21}\\ z_{12} & z_{22}\end{bmatrix}
= \begin{bmatrix}
z_{11}+2z_{12} & z_{21}+2z_{22}\\
z_{12} & z_{22}\\
-z_{11}+z_{12} & -z_{21}+z_{22}
\end{bmatrix}.
$$

它验证了(3.2.11) 是正确的.

(4) 设$\rho: M_2 \to M_{2\times 4}$, 使得$Z \mapsto Z^{\mathrm{T}}D$, 那么

$$M_\rho = (D^{\mathrm{T}} \otimes I_2)W_{[2,2]} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

因此

$$\begin{aligned} V_c(Z^{\mathrm{T}}D) &= M_\rho V_c(Z) \\ &= \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \\ z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} z_{11}+z_{21} \\ z_{12}+z_{22} \\ z_{11} \\ z_{12} \\ z_{11}+2z_{21} \\ z_{12}+2z_{22} \\ -z_{11}+z_{21} \\ -z_{12}+z_{22} \end{bmatrix}. \end{aligned} \tag{3.2.12}$$

直接计算有

$$\begin{aligned} Z^{\mathrm{T}}D &= \begin{bmatrix} z_{11} & z_{21} \\ z_{12} & z_{22} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & -1 \\ 1 & 0 & 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} z_{11}+z_{21} & z_{11} & z_{11}+2z_{21} & -z_{11}+z_{21} \\ z_{12}+z_{22} & z_{12} & z_{12}+2z_{22} & -z_{12}+z_{22} \end{bmatrix}. \end{aligned}$$

这验证了(3.2.12) 是正确的.

(5) 设$\rho : M_2 \to M_{3\times 4}$, 使得$Z \mapsto AZB + CZ^{\mathrm{T}}D$. 利用(3.2.8) 得到

$$M_\rho = (B^{\mathrm{T}} \otimes A) + (D^{\mathrm{T}} \otimes C)W_{[2,2]} = \begin{bmatrix} 2 & 0 & 4 & 0 \\ 2 & 1 & 5 & 3 \\ -1 & 0 & 1 & 3 \\ 0 & 1 & 3 & -1 \\ -2 & -1 & 3 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 2 & 3 & 3 \\ 0 & 0 & 3 & 3 \\ -1 & -2 & 1 & 3 \\ 0 & 0 & -2 & 2 \\ 2 & 1 & -1 & 1 \\ 1 & 0 & -1 & 1 \end{bmatrix}.$$

因此

$$\begin{aligned} & V_c(AZB + CZ^{\mathrm{T}}D) \\ = & M_\rho V_c(Z) \\ = & \begin{bmatrix} 2 & 0 & 4 & 0 \\ 2 & 1 & 5 & 3 \\ -1 & 0 & 1 & 3 \\ 0 & 1 & 3 & -1 \\ -2 & -1 & 3 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 2 & 3 & 3 \\ 0 & 0 & 3 & 3 \\ -1 & -2 & 1 & 3 \\ 0 & 0 & -2 & 2 \\ 2 & 1 & -1 & 1 \\ 1 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \\ z_{12} \\ z_{22} \end{bmatrix} = \begin{bmatrix} 2z_{11} + 4z_{12} \\ 2z_{11} + z_{21} + 5z_{12} + 3z_{22} \\ -z_{11} + z_{12} + 3z_{22} \\ z_{21} + 3z_{12} - z_{22} \\ -2z_{11} - z_{21} + 3z_{12} + z_{22} \\ -z_{11} - z_{21} + z_{12} + z_{22} \\ z_{11} + 2z_{21} + 3z_{12} + 3z_{22} \\ 3z_{12} + 3z_{22} \\ -z_{11} - 2z_{21} + z_{12} + 3z_{22} \\ -2z_{12} + 2z_{22} \\ 2z_{11} + z_{21} - z_{12} + z_{22} \\ z_{11} - z_{12} + z_{22} \end{bmatrix}. \end{aligned} \tag{3.2.13}$$

直接计算得

$$
AZB + CZ^{\mathrm{T}}D = \begin{bmatrix} 2z_{11}+4z_{12} & z_{21}+3z_{12}-z_{22} & z_{11}+2z_{21}+3z_{12}+3z_{22} & -2z_{12}+2z_{22} \\ 2z_{11}+z_{21}+5z_{12}+3z_{22} & -2z_{11}-z_{21}+3z_{12}+z_{22} & 3z_{12}+3z_{22} & 2z_{11}+z_{21}-z_{12}+z_{22} \\ -z_{11}+z_{12}+3z_{22} & -z_{11}-z_{21}+z_{12}+z_{22} & -z_{11}-2z_{21}+z_{12}+3z_{22} & z_{11}-z_{12}+z_{22} \end{bmatrix}.
$$

这验证(3.2.13) 是正确的. □

映射(3.2.7) 由两种不同类型的项组成. 它可以用于任意有限项的情形. 一种特别有用的情况是所有矩阵均为同一维数的方阵. 下面给一个数值例子来说明.

例 3.2.5 设

$$
A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix},
$$

$$
C = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}.
$$

考虑

$$
L(Z) = AZ + ZB + CZ^{\mathrm{T}}D + AZB. \tag{3.2.14}
$$

利用(3.2.8) 及表3.2.1 可得映射的矩阵表达式为

$$
\begin{aligned}
M_L^c &= I_2 \otimes A + B^{\mathrm{T}} \otimes I_2 + (D^{\mathrm{T}} \otimes C)W_{[2]} + B^{\mathrm{T}} \otimes A \\
&= \begin{bmatrix} 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & 2 \end{bmatrix} \\
&\quad + \begin{bmatrix} 1 & 0 & -2 & 0 \\ 1 & -1 & -2 & 2 \\ 2 & 0 & 1 & 0 \\ 2 & -2 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ -1 & -1 & 2 & 2 \\ 1 & -1 & -2 & 2 \end{bmatrix}
\end{aligned}
$$

$$
= \begin{bmatrix} 2 & -1 & 2 & 1 \\ 0 & -1 & -2 & 4 \\ 0 & 0 & 5 & 3 \\ 3 & -2 & 5 & 4 \end{bmatrix}.
$$

□

为了方便起见, 我们在表3.2.2中列出了一些有用的矩阵线性映射的矩阵表示.

表 3.2.2　矩阵线性映射的列展开表示

映射名称	记号	ρ_A	$M^c_{\rho_A}$
Lyapunov 映射	L_A	$Z \mapsto AZ + ZA^{\mathrm{T}}$	$I \otimes A + A \otimes I$
一般 Lyapunov 映射	L_{AB}	$Z \mapsto AZ + ZB$	$I \otimes A + B^{\mathrm{T}} \otimes I$
辛映射	S_A	$Z \mapsto AZ + Z^{\mathrm{T}}A$	$I \otimes A + (A^{\mathrm{T}} \otimes I)W$
伴随映射	ad_A	$Z \mapsto AZ - ZA$	$I \otimes A - A^{\mathrm{T}} \otimes I$
共轭映射	C^A_j	$Z \mapsto AZA^{-1}$	$A^{-\mathrm{T}} \otimes A$
相合映射	C^A_g	$Z \mapsto AZA^{\mathrm{T}}$	$A \otimes A$

利用上面的矩阵表示, 我们可以得到一些有用的公式.

命题 3.2.6　*设*$A \in M_{m\times n}$, $B \in M_{p\times q}$, *那么*

$$
(I_p \otimes A)W_{[n,p]} = W_{[m,p]}(A \otimes I_p), \tag{3.2.15}
$$

$$
W_{[m,p]}(A \otimes B)W_{[q,n]} = (B \otimes A). \tag{3.2.16}
$$

证明　设$Z \in M_{p\times n}$, 考虑表示$Z \mapsto AZ^{\mathrm{T}}$, 它可以通过以下两种方式实现:

(1) $Z \mapsto Z^{\mathrm{T}} \mapsto AZ^{\mathrm{T}}$: 注意到首先是用$W_{[n,p]}$ 作用, 然后是$I_p \otimes A$. 因此它的矩阵表示是: $(I_p \otimes A)W_{[n,p]}$.

(2) $Z \mapsto ZA^{\mathrm{T}} \mapsto (ZA^{\mathrm{T}})^{\mathrm{T}} = AZ^{\mathrm{T}}$: 先用$A \otimes I_p$, 再用$W_{[m,p]}$. 因此, 同样的映射也可以被表示成: $W_{[m,p]}(A \otimes I_p)$.

于是(3.2.15) 成立.

对于(3.2.16), 设$Z \in M_{q\times n}$, 考虑表示$Z \mapsto AZ^{\mathrm{T}}B^{\mathrm{T}}$, 我们也可以通过以下两种方式实现:

(1) $Z \mapsto ZA^{\mathrm{T}} \mapsto BZA^{\mathrm{T}} \mapsto (BZA^{\mathrm{T}})^{\mathrm{T}}$: 它是由$W_{[m,p]}(I_m \otimes B)(A \otimes I_q)$ 实现的, 也就是$W_{[m,p]}(A \otimes B)$.

(2) $Z \mapsto Z^{\mathrm{T}} \mapsto AZ^{\mathrm{T}} \mapsto AZ^{\mathrm{T}}B^{\mathrm{T}}$: 它由$(B \otimes I_m)(I_q \otimes A)W_{[n,q]}$ 实现, 就是$(B \otimes A)W_{[n,q]}$.

于是有

$$W_{[m,p]}(A \otimes B) = (B \otimes A)W_{[n,q]}.$$

用$W_{[q,n]}$ 右乘等式两边就得到(3.2.16). □

可以看出(3.2.15) 也可由(3.2.16) 得到, 后面我们将会用到它们.

作为另一个应用, 我们计算一般线性群的李代数, 称为一般线性代数. 一般线性群, 记作$GL(n,\mathbb{R})$, 指的是所有$n \times n$ 可逆实矩阵在普通矩阵乘法下形成的群. 它又可以看作$\mathbb{R}^{n\times n}$ 的一个开子集, 因此也是个解析流形, 而且, 在这个流形结构下, 乘法和逆这两个运算都是解析的, 因此, 它是一个李群. 李群的所有左不变向量场构成它的李代数.

例 3.2.7 考虑一般线性群$GL(n,\mathbb{R})$. 设F 是由$F(I) = A$ 生成的左不变向量场. 将这个$n \times n$ 矩阵A 看成切空间的向量, 写成典型的局部坐标表示形式, 即将每个元素看成一个坐标分量, 则得$V_c(A)$ (当然也可以用$V_r(A)$). 记左平移为$\phi_X^L : P \mapsto XP, \forall P \in GL(n,\mathbb{R})$, 当将$X$ 看作$\mathbb{R}^{n\times n}$的一个点时, 由(3.2.14) 知ϕ_X^L 的列展开应为$I_n \otimes X$. 因此$\phi_X^L(P) = (I_n \otimes X)p$, 其中$p = V^c(P)$. 令$a = V^c(A)$, 现在将左平移映射看作$GL(n,\mathbb{R})$ 上的一个微分同胚, 它可将I 点的切向量A 移到X 点, 其值为

$$\frac{\partial \phi_X^L(P)}{\partial p} a = (I_n \otimes X)a.$$

但因为$F(X)$ 是左不变的, A 左平移到X 的向量正好是F 在该点的值, 故

$$V^c(F(X)) = J_{\phi_X^L} a = \frac{\partial (I_n \otimes X)x}{\partial x} a = (I_n \otimes X)a = V^c(XA).$$

最后一个等号由(3.2.14) 得到. 因此, $F(X)$ 在X 的矩阵形式是XA.

接着, 给定两个左不变向量场F 和W, 分别由A 和B 生成. 由上面可知, F 和W 的矩阵形式分别是$F(X) = XA$ 和$W(X) = XB$. 利用(3.2.14), 在向量形式下它们可以分别表示成

$$F(x) = (A^{\mathrm{T}} \otimes I_n)x, \quad W(x) = (B^{\mathrm{T}} \otimes I_n)x.$$

根据公式[68]

$$[F(x), W(x)] = \frac{\partial W}{\partial x} F - \frac{\partial F}{\partial x} W,$$

我们有

$$
\begin{aligned}
[F(x), W(x)] &= (B^{\mathrm{T}} \otimes I_n)(A^{\mathrm{T}} \otimes I_n)x - (A^{\mathrm{T}} \otimes I_n)(B^{\mathrm{T}} \otimes I_n)x \\
&= ((B^{\mathrm{T}} \otimes I_n)(A^{\mathrm{T}} \otimes I_n) - (A^{\mathrm{T}} \otimes I_n)(B^{\mathrm{T}} \otimes I_n))x \\
&= (B^{\mathrm{T}}A^{\mathrm{T}} \otimes I_n - A^{\mathrm{T}}B^{\mathrm{T}} \otimes I_n)x = ((AB - BA) \otimes I_n)x.
\end{aligned}
$$

利用(3.2.14), 有$[F(x), W(x)]$ 的矩阵表示是$(AB - BA)X$, 也就是由$AB - BA$ 生成的左不变向量场. 因此, 李群$GL(n, \mathbb{R})$ 的李代数$gl(n, \mathbb{R})$ 上的李括号是

$$
[A, B] = AB - BA. \tag{3.2.17}
$$

□

下面的命题给出了线性映射的行展开, 实际上它也可以通过它的列展开得到.

命题 3.2.8 (3.2.7) 的行展开是

$$
M^r = (A \otimes B^{\mathrm{T}}) + (C \otimes D^{\mathrm{T}})W_{[n,p]}. \tag{3.2.18}
$$

证明 这里也先列出四种基本线性映射的行展开, 当然这些表示本身也很有用.

我们逐个证明它们. 利用(3.1.10) 和(3.1.2), 我们有

$$
V_r(AZ) = AV_r(Z) = (A \otimes I_p)V_r(Z),
$$

这就证明了第一个. 对于第二个, 有

$$
V_r(ZB) = V_c(B^{\mathrm{T}}Z^{\mathrm{T}}) = (I_n \otimes B^{\mathrm{T}})V_c(Z^{\mathrm{T}}) = (I_n \otimes B^{\mathrm{T}})V_r(Z).
$$

而对于第三个, 我们有

$$
\begin{aligned}
V_r(CZ^{\mathrm{T}}) &= V_c(ZC^{\mathrm{T}}) = (C \otimes I_n)V_c(Z) \\
&= (C \otimes I_n)W_{[n,p]}W_{[p,n]}V_c(Z) \\
&= (C \otimes I_n)W_{[n,p]}V_r(Z).
\end{aligned}
$$

然后, 由

$$
\begin{aligned}
V_r(Z^{\mathrm{T}}D) &= V_c(D^{\mathrm{T}}Z) = (I_p \otimes D^{\mathrm{T}})V_c(Z) \\
&= (I_p \otimes D^{\mathrm{T}})W_{[n,p]}W_{[p,n]}V_c(Z) \\
&= (I_p \otimes D^{\mathrm{T}})W_{[n,p]}V_r(Z).
\end{aligned}
$$

就证明了最后一个. 利用表3.2.3,

表 3.2.3 行展开

ρ:	M_ρ^r
$Z \mapsto AZ$	$A \otimes I_p$
$Z \mapsto ZB$	$I_n \otimes B^{\mathrm{T}}$
$Z \mapsto CZ^{\mathrm{T}}$	$(C \otimes I_n)W_{[n,p]}$
$Z \mapsto Z^{\mathrm{T}}D$	$(I_p \otimes D^{\mathrm{T}})W_{[n,p]}$

我们得到复合映射

$$\begin{aligned} &V_r(AZB + CZ^{\mathrm{T}}D) \\ &= [(I_m \otimes B^{\mathrm{T}})(A \otimes I_p) + (I_m \otimes D^{\mathrm{T}})(C \otimes I_n)W_{[n,p]}]V_r(Z) \\ &= [A \otimes B^{\mathrm{T}} + (C \otimes D^{\mathrm{T}})W_{[n,p]}]V_r(Z). \end{aligned} \tag{3.2.19}$$

于是(3.2.18) 成立. □

作为应用, 我们考虑Hautus 方程. 设$A_i \in M_{n\times m}$, $q_i(t)$, $i = 1,\cdots,k$ 是一些多项式, $S \in M_{p\times p}$, $R \in M_{n\times p}$, $X \in M_{m\times p}$. 称下面的关于未知量X 的矩阵方程是Hautus 方程:

$$A_1Xq_1(S) + \cdots + A_kXq_k(S) = R. \tag{3.2.20}$$

为了说明Hautus 方程的重要性, 我们考虑它的一些特殊情形.

设$A \in M_{n\times n}$, $S \in M_{p\times p}$, $R \in M_{n\times p}$. 称下面的方程是Sylvester 方程, 它在控制理论中很重要:

$$AX - XS = R. \tag{3.2.21}$$

可以看出, 它是(3.2.20) 在$A_1 = A$, $A_2 = I$, $q_1(t) = 1$, $q_2(t) = -t$ 时的特殊情形.

设A 和S 都是方阵, 并且B, P, C, Q 都是具有合适维数的矩阵. 称下面的方程是调节方程, 它是在研究控制系统输出调节问题时遇到的.

$$\begin{aligned} \Pi S &= A\Pi + B\Gamma + P, \\ 0 &= C\Pi + Q. \end{aligned} \tag{3.2.22}$$

(3.2.22) 可以转化为一个Hautus 方程:

$$A_1X - A_2XS = R,$$

其中

$$A_1 = \begin{bmatrix} A & B \\ C & 0 \end{bmatrix}, A_2 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, R = \begin{bmatrix} -P \\ -Q \end{bmatrix}, X = \begin{bmatrix} \Pi \\ \Gamma \end{bmatrix}.$$

定理 3.2.9 Hautus 方程对于每个R 都有解, 当且仅当矩阵

$$A(\lambda) = A_1 q_1(\lambda) + \cdots + A_k q_k(\lambda) \tag{3.2.23}$$

的n 个行对于S 的每个特征值, 即对每一个$\lambda \in \sigma(S)$, 均是线性无关的. 而且, 如果$n = m$, 则解是唯一的.

证明 设$T \in O(p, \mathbb{R})$, 且令

$$\tilde{X} = XT, \quad \tilde{S} = T^{-1}ST, \quad \tilde{R} = R.$$

则方程(3.2.20) 可以转化为

$$A_1 \tilde{X} q_1(\tilde{S}) + \cdots + A_k \tilde{X} q_k(\tilde{S}) = \tilde{R}. \tag{3.2.24}$$

显然, 对于每个R, 都存在解X, 与对于$\tilde{R}$ 存在解$\tilde{X}$等价, 并且有$X = \tilde{X}T^{-1}$. 因此, 不失一般性, 我们可以假设, S 就是它的Jordan 标准型.

利用命题3.2.3, (3.2.20) 可以转化为

$$\left[q_1(S^{\mathrm{T}}) \otimes A_1 + \cdots + q_k(S^{\mathrm{T}}) \otimes A_k\right] x = r, \tag{3.2.25}$$

其中$x = V_c(X)$, $r = V_c(R)$. 由于S 具有Jordan 标准型

$$S = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \lambda_p \end{bmatrix},$$

于是(3.2.25) 可以被表示成$Ex = r$, 这里

$$E = \begin{bmatrix} Q(\lambda_1) & 0 & \cdots & 0 \\ * & Q(\lambda_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ * & \cdots & * & Q(\lambda_p) \end{bmatrix},$$

其中

$$Q(t) = q_1(t)A_1 + \cdots + q_k(t)A_k.$$

于是立即就可以看出结论成立. □

推论 3.2.10 Sylvester 方程(3.2.21) 对于每个R 都有解, 当且仅当A 和S 没有相同的特征值. 而且, 这时的解是唯一的.

证明 对于Sylvester 方程, 我们有

$$A(\lambda) = A - I\lambda.$$

对于任意的$\lambda \in \sigma(S)$, $A(\lambda)$ 非异等价于A 和S 没有相同的特征值. □

3.3 矩阵的形式转换

本节我们将考虑将一种矩阵转化成另一种或者将一种矩阵表示转化成另一种的问题. 其中后者尤为重要, 因为我们可以通过它将一般的矩阵表示转化成标准的多项式形式.

首先, 我们考虑如何将一个变量为矩阵X 的多项式转化为关于变量展开.

设$A \in M_{m\times n}$, 记A_i 为它的第i 列. 定义两个映射

$$\pi_s^r, \pi_s^c : M_{m\times n} \to M_{m\times ns^2},$$

分别为

$$\pi_s^r(A) = A[I_n \otimes V_r^{\mathrm{T}}(I_s)], \tag{3.3.1}$$

和

$$\pi_s^c(A) = A[I_n \otimes \delta_s^1 \ I_n \otimes \delta_s^2 \ \cdots \ I_n \otimes \delta_s^s], \tag{3.3.2}$$

其中δ_s^i 表示I_s 的第i 行. 这两个映射有如下关系

命题 3.3.1

$$\pi_s^c(A) = \pi_s^r(A)W_{[s,n]}, \quad \forall A \in M_{m\times n}. \tag{3.3.3}$$

证明 仔细计算有

$$\begin{aligned}\pi_s^r(A) = & \Big[\underbrace{\underbrace{A_1\, 0 \cdots 0}_{s}\ \underbrace{0\, A_1\, 0 \cdots 0}_{s} \cdots \underbrace{0 \cdots A_1}_{s}}_{s} \cdots \\ & \underbrace{\underbrace{A_n\, 0 \cdots 0}_{s}\ \underbrace{0\, A_n\, 0 \cdots 0}_{s} \cdots \underbrace{0 \cdots A_n}_{s}}_{s}\Big],\end{aligned} \tag{3.3.4}$$

和

$$\pi_s^c(A) = \Big[\underbrace{\underbrace{A_1\,0\,\cdots\,0}_{s}\ \underbrace{A_2\,0\,\cdots\,0}_{s}\ \cdots\ \underbrace{A_n\,0\,\cdots\,0}_{s}}_{n}\cdots$$
$$\underbrace{\underbrace{0\,\cdots\,A_1}_{s}\ \underbrace{0\,\cdots\,A_2}_{s}\ \cdots\ \underbrace{0\,\cdots\,A_n}_{s}}_{n}\Big]. \tag{3.3.5}$$

记

$$H_{ij} = (\underbrace{0\cdots0}_{j-1}A_i\underbrace{0\cdots0}_{s-j}),\quad i=1,\cdots,n;\ j=1,\cdots,s.$$

则得

$$\pi_s^r(A) = \Big[H_{11}H_{12}\cdots H_{1s}\ \cdots\ H_{n1}H_{n2}\cdots H_{ns}\Big];$$
$$\pi_s^c(A) = \Big[H_{11}H_{21}\cdots H_{n1}\ \cdots\ H_{1s}H_{2s}\cdots H_{ns}\Big].$$

即在 (3.3.4) 中 $\{H_{ij}\}$ 依索引 $Id(i,j;n,s)$ 排列, 而在 (3.3.4) 中 $\{H_{ij}\}$ 依索引 $Id(j,i;s,n)$ 排列. 根据命题 3.1.4 立即可知 (3.3.3) 成立. □

现在, 我们准备将一个常值矩阵和一个变量矩阵的乘积转化为变量展开.

命题 3.3.2　设 $A\in M_{m\times n}$, $X\in M_{n\times s}$, 那么

$$AX = \pi_s^r(A)V_r(X),\quad 或\quad AX = \pi_s^r(A)W_{[s,n]}V_c(X); \tag{3.3.6}$$

或者

$$AX = \pi_s^c(A)V_c(X),\quad 或\quad AX = \pi_s^c(A)W_{[n,s]}V_r(X). \tag{3.3.7}$$

证明　利用(3.3.4), 直接计算得到

$$\pi_s^r(A)V_r(X) = \left[\sum_{k=1}^n A_kx_{1k}, \sum_{k=1}^n A_kx_{2k}, \cdots, \sum_{k=1}^n A_kx_{nk}\right] = AX,$$

其中A_k 表示A 的第k 列. 这就证明了(3.3.6) 中的第一个等式.

(3.3.6) 中的第二个等式可以由第一个等式和命题1.5.4 得到. (3.3.7) 的证明类似. □

结合上面的这些公式和矩阵线性映射的一些结果, 我们可以得到一些有用的公式.

命题 3.3.3 设$X \in M_{m\times n}$, $A \in M_{n\times s}$, 那么

$$XA = (I_m \otimes V_r^{\mathrm{T}}(I_s))W_{s,m}A^{\mathrm{T}}V_c(X). \tag{3.3.8}$$

证明 利用(3.3.7) 和(3.2.14), 我们有

$$\begin{aligned} XA &= I_m(XA) = \pi_s^c(I_m)V_c(XA) \\ &= \pi_s^c(I_m)(A^{\mathrm{T}} \otimes I_m)V_c(X). \end{aligned} \tag{3.3.9}$$

利用(3.3.1) 和(3.3.3), 有

$$\pi_s^c(I_m) = \left(I_m \otimes V_r^{\mathrm{T}}(I_s)\right) \otimes W_{[s,m]}.$$

再由(3.1.4) 得到

$$(A^{\mathrm{T}} \otimes I_m)V_c(X) = A^{\mathrm{T}}V_c(X).$$

将它们代入(3.3.9) 就得到(3.3.8). □

下面给出一个例子说明如何将一个矩阵多项式表示成它的元素的多项式.

例 3.3.4 设$A, B, C, Z \in M_n$, 考虑映射$Z \mapsto AZBZC$. 我们想将它表示成它的元素的二次型.

利用(3.2.8) 和(3.3.8), 我们有

$$\begin{aligned} V_c(AZBZC) &= (C^{\mathrm{T}} \otimes A)V_c(ZBZ) \\ &= (C^{\mathrm{T}} \otimes A) \ltimes (BZ)^{\mathrm{T}} \ltimes V_c(Z). \end{aligned} \tag{3.3.10}$$

再应用(3.3.20) 或者应用(3.2.16)和(3.1.4) 于$(BZ)^{\mathrm{T}}$ 得到

$$\begin{aligned} (BZ)^{\mathrm{T}} &= Z^{\mathrm{T}}B^{\mathrm{T}} = \left(I_n \otimes V_r^{\mathrm{T}}(I_n)\right) W_{[n]} \ltimes B \ltimes V_c(Z^{\mathrm{T}}) \\ &= \left(I_n \otimes V_r^{\mathrm{T}}(I_n)\right) W_{[n]}(B \otimes I_n)W_{[n]}V_r(Z^{\mathrm{T}}) \\ &= \left(I_n \otimes V_r^{\mathrm{T}}(I_n)\right) (I_n \otimes B)V_c(Z). \end{aligned}$$

最后, 我们有

$$V_c(AZBZC) = (C^{\mathrm{T}} \otimes A) \left(I_n \otimes V_r^{\mathrm{T}}(I_n)\right) (I_n \otimes B)V_c^2(Z). \tag{3.3.11}$$

□

利用上面的表示我们可以很方便地研究矩阵函数的一些性质. 下面给出一个例子, 在后面也将会用到它.

例 3.3.5 考虑一个线性控制系统

$$\dot{x} = Ax + Bu, \quad x \in \mathbb{R}^n,\ u \in \mathbb{R}^m.$$

当使用一个状态反馈

$$u = Fx + v,$$

并且考虑解耦问题时, 为了计算解耦矩阵, 下面的形式很重要[68]:

$$(A + BF)^k. \tag{3.3.12}$$

我们将$(A+BF)^k$ 看作F 的函数来研究它的性质. 半张量积可以将它按如下方式表示成一个F 的多项式型函数: 首先, 利用命题3.3.2, 我们可以将$A+BF$ 转化为

$$A + BF = A + Ef,$$

其中$f = V_r(F)$, $E = \pi_m^r(B)$, 这也可以由(3.3.1) 得到. 于是有

$$\begin{aligned}(A + E \ltimes f)^2 &= (A + E \ltimes f)(A + E \ltimes f)\\ &= A^2 + [A \ltimes E + E \ltimes (I_n \otimes A)] \ltimes f + E \ltimes (I_n \otimes E) \ltimes f^2.\end{aligned}$$

一般我们可以假设

$$(A + BF)^k = A^k + P_1^k \ltimes f + P_2^k \ltimes f^2 + \cdots + P_k^k \ltimes f^k, \tag{3.3.13}$$

那么

$$\begin{aligned}(A + BF)^{k+1} &= (A + E \ltimes f)(A^k + P_1^k \ltimes f + P_2^k \ltimes f^2 + \cdots + P_k^k \ltimes f^k)\\ &= A^{k+1} + [A \ltimes P_1^k + E \ltimes (I_n \otimes A^k)] \ltimes f\\ &\quad +[A \ltimes P_2^k + E \ltimes (I_n \otimes P_1^k)] \ltimes f^2\\ &\quad + \cdots + [A \ltimes P_k^k + E \ltimes (I_n \otimes P_k^k)] \ltimes f^{k+1}.\end{aligned}$$

为了符号的方便, 我们记$P_0^k := A^k$, 那么对于(3.3.13) 中的系数, 我们有一个递推式

$$\begin{cases} P_1^1 = E, \\ P_i^{k+1} = A \ltimes P_i^k + E \ltimes (I_n \otimes P_{i-1}^k), \quad i = 1, 2, \cdots, k+1. \end{cases} \tag{3.3.14}$$

注意到$P_i^k \in M_{n \times n^{i+1}}$, 并且很容易计算它. 实际上, (3.3.12) 中的矩阵的每一个元素都是$\{f_{ij}\}$ 的函数, 但是在这种形式下难以计算, 而现在(3.3.13)~(3.3.14) 就提供了计算每个元素的函数的简便方法. □

粗略的说, 换位矩阵也可以将一个矩阵转化为一个向量. 在某些情况下使用下面的公式将会很方便.

命题 3.3.6 (1) 设Z 是一个t 维行向量, $A \in M_{m\times n}$, 那么

$$ZW_{[m,t]}A = AZW_{[n,t]} = A \otimes Z. \tag{3.3.15}$$

(2) 设Y 是一个t 维列向量, A 同上, 那么

$$AW_{[t,n]}Y = W_{[t,m]}YA = A \otimes Y. \tag{3.3.16}$$

证明 (1) 利用(1.5.7), 直接计算有

$$ZW_{[m,t]} = \sum_{j=1}^{t} z_j I_m \otimes \delta^j = I_m \otimes Z. \tag{3.3.17}$$

利用(3.1.3), 我们有

$$ZW_{[m,t]}A = (I_m \otimes Z)A = (I_m \otimes Z)(A \otimes I_t) = A \otimes Z.$$

同理, 我们有

$$AZW_{[n,t]} = A(I_n \otimes Z) = (A \otimes I_1)(I_n \otimes Z) = A \otimes Z.$$

(2) 在 (3.3.15) 中用 A^{T} 替换 A, Y^{T} 替换 Z, 然后两边同时作转置. 注意到 $W_{[m,n]}^{\mathrm{T}} = W_{[n,m]}$, 于是 (3.3.16) 成立. □

下面我们证明一个有用的引理.

引理 3.3.7 设$A \in M_{m\times n}$, 那么

$$W_{[m,q]} \ltimes A \ltimes W_{[q,n]} = I_q \otimes A. \tag{3.3.18}$$

证明 对于$X \in M_{n\times q}$, 根据(3.1.10)我们有

$$V_r(AX) = A \ltimes V_r(X) = A \ltimes W_{[q,n]}V_c(X).$$

两边左乘$W_{[m,q]}$ 得到

$$V_c(AX) = (W_{[m,q]} \ltimes A \ltimes W_{[q,n]})V_c(X). \tag{3.3.19}$$

比较(3.2.14) 的第一式和(3.3.19), 就得到(3.3.18). □

左半张量积不满足交换律, 但是有时我们需要交换因子的次序. 下面的公式, 在一定意义上可以看作是"交换次序"公式, 它在多元多项式计算中非常有用.

命题 3.3.8 给定矩阵 $A \in \mathcal{M}_{m\times n}$.

(1) 设 $Z \in \mathbb{R}^t$ 是一个列向量, 则

$$AZ^{\mathrm{T}} = Z^{\mathrm{T}} W_{[m,t]} A W_{[t,n]} = Z^{\mathrm{T}}(I_t \otimes A). \tag{3.3.20}$$

(2) 设 $Z \in \mathbb{R}^t$ 是一个列向量, 则

$$ZA = W_{[m,t]} A W_{[t,n]} Z = (I_t \otimes A)Z. \tag{3.3.21}$$

(3) 设 $X \in \mathbb{R}^m$ 是一个列向量, 则

$$X^{\mathrm{T}} A = [V_r(A)]^{\mathrm{T}} X. \tag{3.3.22}$$

(4) 设 $Y \in \mathbb{R}^n$ 是一个列向量, 则

$$AY = Y^{\mathrm{T}} V_c(A). \tag{3.3.23}$$

(5) 设 $X \in \mathbb{R}^m$ 是一个列向量, $Y \in \mathbb{R}^n$ 是一个行向量, 则

$$XY = Y W_{[m,n]} X. \tag{3.3.24}$$

证明 将(3.3.15) 中的第一个等号两边同时右乘 $W_{[t,n]}$, 并且利用 $W_{[n,t]}^{-1} = W_{[t,n]}$ 就有(3.3.20) 的第一个等号成立. 由第一个等号, 再利用(3.3.18) 就得到第二个等号.

同理, 将(3.3.16) 中的第一个等号两边同时左乘 $W_{[m,t]}$ 得到(3.3.21) 的第一个等号. 再对第一个等号利用(3.3.18) 就得到第二个等号.

我们将(3.3.22) 和(3.3.23) 的证明留给读者.

注意到 $W_{[1,t]} = W_{[t,1]} = I_t$, 由(3.3.20) 或(3.3.21) 立即可得(3.3.24). □

命题 3.3.9 设 $A \in \mathcal{M}_{m\times n}$, $B \in \mathcal{M}_{s\times t}$, 那么

$$A \otimes B = W_{[s,m]} \ltimes B \ltimes W_{[m,t]} \ltimes A = (I_m \otimes B) \ltimes A. \tag{3.3.25}$$

证明 记B 的第i 行为B^i, 通过直接计算有

$$A \otimes B = \begin{bmatrix} a_{11}B^1 & \cdots & a_{1n}B^1 \\ a_{11}B^2 & \cdots & a_{1n}B^2 \\ \vdots & & \vdots \\ a_{11}B^s & \cdots & a_{1n}B^s \\ \vdots & & \vdots \\ a_{m1}B^1 & \cdots & a_{mn}B^1 \\ a_{m1}B^2 & \cdots & a_{mn}B^2 \\ \vdots & & \vdots \\ a_{m1}B^s & \cdots & a_{mn}B^s \end{bmatrix} = \begin{bmatrix} B^1 \ltimes A^1 \\ B^2 \ltimes A^1 \\ \vdots \\ B^s \ltimes A^1 \\ \vdots \\ B^1 \ltimes A^m \\ B^2 \ltimes A^m \\ \vdots \\ B^s \ltimes A^m \end{bmatrix}. \tag{3.3.26}$$

对B 的每一行, 利用(3.3.15), 有

$$B \ltimes W_{[m,t]} \ltimes A = \begin{bmatrix} A \otimes B^1 \\ A \otimes B^2 \\ \vdots \\ A \otimes B^s \end{bmatrix} = \begin{bmatrix} B^1 \ltimes A^1 \\ B^1 \ltimes A^2 \\ \vdots \\ B^1 \ltimes A^m \\ \vdots \\ B^s \ltimes A^1 \\ B^s \ltimes A^2 \\ \vdots \\ B^s \ltimes A^m \end{bmatrix}. \tag{3.3.27}$$

比较(3.3.26) 和(3.3.27) 可以看出, $B^i \ltimes A^j$ 在(3.3.27) 中是按照索引$Id(i,j;s,m)$ 排列的, 而在(3.3.26) 中则是按照索引$Id(j,i;m,s)$ 排列的. 现在为了将$Id(i,j;s,m)$ 转化为$Id(j,i;m,s)$, 我们需要用$W_{[s,m]}$ 左乘(3.3.27). 这样(3.3.25) 的第一个等号成立.

对(3.3.25) 中的第一个等号应用(3.3.18) 就得到了第二个等号. □

例 3.3.10 设

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix},$$

那么$m = n = 2$, $s = 3$, $t = 2$, 于是

$$W_{[32]} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad W_{[22]} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\begin{aligned} & W_{[32]} \ltimes B \ltimes W_{[22]} \ltimes A \\ = & \begin{bmatrix} b_{11} & b_{12} & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ b_{31} & b_{32} & 0 & 0 \\ 0 & 0 & b_{11} & b_{12} \\ 0 & 0 & b_{21} & b_{22} \\ 0 & 0 & b_{31} & b_{32} \end{bmatrix} \ltimes \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \\ = & \begin{bmatrix} a_{11}b_{11} & a_{11}b_{12} & a_{12}b_{11} & a_{12}b_{12} \\ a_{11}b_{21} & a_{11}b_{22} & a_{12}b_{21} & a_{12}b_{22} \\ a_{11}b_{31} & a_{11}b_{32} & a_{12}b_{31} & a_{12}b_{32} \\ a_{21}b_{11} & a_{21}b_{12} & a_{22}b_{11} & a_{22}b_{12} \\ a_{21}b_{21} & a_{21}b_{22} & a_{22}b_{21} & a_{22}b_{22} \\ a_{21}b_{31} & a_{21}b_{32} & a_{22}b_{31} & a_{22}b_{32} \end{bmatrix} = A \otimes B. \end{aligned}$$

□

下面的推论是一个十分有用的等式.

推论 3.3.11　设$C \in M_{s\times t}$, 则对任一整数$m > 0$ 均有

$$W_{[s,m]} \ltimes C \ltimes W_{[m,t]} = I_m \otimes C. \tag{3.3.28}$$

证明　在(3.3.25) 式中取$B = C$ 及$A = I_m$ 即得. □

最后, 我们考虑如何将一个矩阵表示成展开式, 或者反之.

命题 3.3.12　设$A \in M_{m\times n}$, 那么

$$V_r(A) = A \ltimes V_r(I_n), \tag{3.3.29}$$

$$V_c(A) = W_{[m,n]} \ltimes A \ltimes V_c(I_n). \tag{3.3.30}$$

证明 直接计算就可得到(3.3.29). 应用(1.5.4) 到(3.3.29) 得到(3.3.30). □

反之, 我们也可以从A 的行或列展开得到A.

命题 3.3.13 设$A \in M_{m\times n}$, 那么

$$A = [I_m \otimes V_r^{\mathrm{T}}(I_n)] \ltimes V_r(A) = [I_m \otimes V_r^{\mathrm{T}}(I_n)] \ltimes W_{[n,m]} \ltimes V_c(A). \tag{3.3.31}$$

证明 分别用I_m 和A 替换(3.3.6) 中的A 和X, 再利用(3.3.1) 就得到第一个等号. 对第一个等号应用(1.5.4) 就得到第二个等号. □

例 3.3.14 给定矩阵$A \in M_{3\times 2}$, 且

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}.$$

那么$m = 3$, $n = 2$. 利用(3.3.28), 我们有

$$V_r(I_n) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

$$A \ltimes V_r(I_n) = \begin{bmatrix} \begin{bmatrix} a_{11} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_{12} \end{bmatrix} \\ \begin{bmatrix} a_{21} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_{22} \end{bmatrix} \\ \begin{bmatrix} a_{31} \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_{32} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ a_{31} \\ a_{31} \end{bmatrix} = V_r(A).$$

利用(3.3.30), 我们有(关于$W_{[3,2]}$ 和$W_{[2,3]}$ 的表达式, 可在例1.5.2 中找到)

$$W_{[3,2]} \ltimes A = \begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ a_{21} & 0 & a_{22} & 0 \\ a_{31} & 0 & a_{32} & 0 \\ 0 & a_{11} & 0 & a_{12} \\ 0 & a_{21} & 0 & a_{22} \\ 0 & a_{31} & 0 & a_{32} \end{bmatrix}$$

那么

$$W_{[3,2]} \ltimes A \ltimes V_r(I_n) = \begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ a_{21} & 0 & a_{22} & 0 \\ a_{31} & 0 & a_{32} & 0 \\ 0 & a_{11} & 0 & a_{12} \\ 0 & a_{21} & 0 & a_{22} \\ 0 & a_{31} & 0 & a_{32} \end{bmatrix} \ltimes \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} = V_c(A).$$

接下来, 我们检验(3.3.31). 对于第一部分我们有

$$\begin{aligned} &[I_3 \otimes V_r^{\mathrm{T}}(I_2)] \ltimes V_r(A) = [I_3 \otimes (1\ 0\ 0\ 1)] \ltimes V_r(A) \\ &= \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \ltimes \begin{bmatrix} a_{11} \\ a_{12} \\ a_{21} \\ a_{22} \\ a_{31} \\ a_{31} \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = A. \end{aligned}$$

为了验证(3.3.31) 的第二部分, 我们有

$$\begin{aligned} &[I_3 \otimes (1\ 0\ 0\ 1)] \ltimes W_{[2,3]} \\ &= \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \ltimes \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

于是

$$
\begin{aligned}
&[I_3 \otimes (1\,0\,0\,1)] \ltimes W_{[2,3]} \ltimes V_c(A) \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \ltimes \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} = A.
\end{aligned}
$$

□

3.4 注释与参考

矩阵间的线性映射是大家都比较熟悉的, 它们在线性控制系统理论中起着关键作用. 相关的参考书很多, 例如文献[8], [65] 等.

直观地说, 李群是这样一种集合, 从几何结构看, 它是一个解析微分流形, 从代数结构看, 它是一个群. 而这两种结构有相容性, 即群运算依几何结构也是解析的. 有人说, 李群是20 世纪发现的最漂亮的数学结构. 这方面的专著很多, 较通俗易懂的读本如文献[10].

Lyapunov 映射与线性系统的稳定性关系密切. 这方面有许多讨论, 例如, 文献[18] 讨论了它的许多性质. 它与切换线性系统的关系, 可见文献[24]. 记S 为$n\times n$ 对称矩阵集合, K 为$n \times n$ 反对称矩阵集合. 显然, 它们都是$M_{n\times n}$ 的子空间. 并且, 容易证明, 它们都是Lyapunov 映射L_A 的不变子空间. 那么, 可以将L_A 分别限制在这两个不变子空间上, 记作L_A^S 及L_A^K.

文献[18] 证明了$\|L_A\| = \max\{\|L_A^S\|, \|L_A^K\|\}$. 在这个基础上文献[18] 提出一个猜想

$$\|L_A\| = \|L_A^S\|.$$

大量的数值计算支持这一结果, 但这个猜想至今尚未能被证明或否定.

Hautus 方程和Sylvester 方程在控制理论中有许多应用. 特别是在输出调节问题中. 关于Hautus 方程及Sylvester 方程的讨论可见文献[72]. 对输出调节有兴趣的读者可参阅文献[64].

习 题 三

3.1 设$A \prec_t B$, 证明

$$\mathrm{rank}(AB) \le \min(t \times \mathrm{rank}(A), \mathrm{rank}(B)).$$

3.2　设$A \succ_t B$, 证明

$$\mathrm{rank}(AB) \le \min(\mathrm{rank}(A), t \times \mathrm{rank}(B)).$$

3.3　设$A \in M_{m\times n}$, $X \in M_{n\times q}$, $Y \in M_{p\times m}$, 利用命题3.1.6 及命题1.5.4, 证明以下两个公式:

$$V_c(AX) = W_{[m,q]} A W_{[q,n]} V_c(X);$$

$$V_r(YA) = W_{[n,p]} A^{\mathrm{T}} W_{[p,m]} V_r(Y).$$

3.4　证明推论3.1.7.

3.5　回顾例 3.2.7, 如果我们使用李代数 $gl(n,\mathbb{R})$ 的右不变向量场, 那么证明 $gl(n,\mathbb{R})$ 上的乘法就是

$$[A, B] = BA - AB.$$

3.6　证明(3.3.22) 和(3.3.23).

3.7　考虑代数Riccati 方程

$$A^{\mathrm{T}}X + XA + Q + XRX = 0,$$

其中A, Q, R 都是$n \times n$ 矩阵. 请将它表示成下面的形式

$$C_0x^2 + C_1x + C_2 = 0,$$

其中$x = V_c(X)$.

3.8　给出一个矩阵方程

$$AXBX^{\mathrm{T}}CX = XDX^{\mathrm{T}},$$

其中A, B, C, D, X 都是$n \times n$ 矩阵. 请将它表示成

$$Fx^3 - Gx^2 = (Fx - G)x^2 = 0$$

的形式, 其中(1) $x = V_r(X)$, (2) $x = V_c(X)$. (从中可以看出在解矩阵方程时后一种表示的优势)

第 4 章　一般半张量积

本章主要考虑矩阵半张量积的一些其他类型. 首先, 我们考虑右半张量积, 它和矩阵的左半张量积类似, 但是它缺少分块乘法的性质, 因此, 没有左半张量积的广泛适用性. 然后, 将左半张量积和右半张量积推广到更一般的情形, 使得具有任意维数的两个矩阵都可作半张量积运算. 最后, 我们考虑一种新的代数, 它将半张量积作为它的乘法.

4.1　右半张量积

首先我们回顾矩阵的张量积 (Kronecker 积). 由 (2.1.13) 容易看出, 对于 $A \in M_{m\times n}$ 和 $B \in M_{p\times q}$, 它们的张量积是

$$A \otimes B = (A \otimes I_p)(I_n \otimes B). \tag{4.1.1}$$

根据命题3.1.2, 左半张量积的一个等价定义是

$$A \ltimes B = \begin{cases} (A \otimes I_t)B, & A \prec_t B, \\ A(B \otimes I_t), & A \succ_t B, \end{cases} \tag{4.1.2}$$

它很像是对(4.1.1) 作左型匹配, 这就是我们为什么称之为左半张量积.

显然, 如果两个矩阵满足左半张量积所要求的倍维数条件, 那么我们也可以作右型匹配, 于是很自然地就有下面的定义.

定义 4.1.1　给出两个矩阵A, B. 假设$A \prec_t B$ 或者$A \succ_t B$, 于是我们定义A和B 的右半张量积为

$$A \rtimes B = \begin{cases} (I_t \otimes A)B, & A \prec_t B, \\ A(I_t \otimes B), & A \succ_t B. \end{cases} \tag{4.1.3}$$

左半张量积的很多性质对于右半张量积仍然成立. 我们不加证明地给出下面的命题.

命题 4.1.2　(1) 右半张量积具有结合律:

$$(A \rtimes B) \rtimes C = A \rtimes (B \rtimes C), \tag{4.1.4}$$

和分配律:

$$(A + B) \rtimes C = A \rtimes C + B \rtimes C, \quad C \rtimes (A + B) = C \rtimes A + C \rtimes B. \tag{4.1.5}$$

(2) 设X, Y 是两个列向量, 那么

$$X \rtimes Y = Y \otimes X. \tag{4.1.6}$$

如果X, Y 是两个行向量, 那么

$$X \rtimes Y = X \otimes Y. \tag{4.1.7}$$

(3)
$$(A \rtimes B)^{\mathrm{T}} = B^{\mathrm{T}} \rtimes A^{\mathrm{T}}. \tag{4.1.8}$$

(4) 如果$M \in M_{m\times pn}$, 那么

$$M \rtimes I_n = M; \tag{4.1.9}$$

如果$M \in M_{m\times n}$, 那么

$$M \rtimes I_{pn} = I_p \otimes M; \tag{4.1.10}$$

如果$M \in M_{pm\times n}$, 那么

$$I_p \rtimes M = M; \tag{4.1.11}$$

如果$M \in M_{m\times n}$, 那么

$$I_{pm} \rtimes M = I_p \otimes M. \tag{4.1.12}$$

在下面的5~10 中, 设A, B 是使右半张量积有定义的具有合适维数的两个方阵.

(5) $A \rtimes B$ 和$B \rtimes A$ 具有相同的特征函数.

(6)
$$\mathrm{tr}(A \rtimes B) = \mathrm{tr}(B \rtimes A). \tag{4.1.13}$$

(7) 如果A 和B 都是正交阵(上三角阵, 下三角阵或对角阵), 那么$A \rtimes B$ 也是.

(8) 如果A 或B 可逆, 那么$A \rtimes B \sim B \rtimes A$.

(9) 如果A 和B 都可逆, 那么

$$(A \rtimes B)^{-1} = B^{-1} \rtimes A^{-1}. \tag{4.1.14}$$

(10) 如果$A \prec_t B$, 那么

$$\det(A \rtimes B) = [\det(A)]^t \det(B); \tag{4.1.15}$$

如果$A \succ_t B$, 那么

$$\det(A \rtimes B) = \det(A)[\det(B)]^t. \tag{4.1.16}$$

一个很自然的问题是: 我们可不可以像在定义2.3.5 中定义左半张量积那样通过“行乘列” 的形式来定义矩阵的右半张量积? 实际上, 我们可以先定义一个右半张量内积:

定义 4.1.3 设$X = (x_1, \cdots, x_s)$ 是一个行向量, $Y = (y_1, \cdots, y_t)^{\mathrm{T}}$ 是一个列向量.

情形1: 如果t 是s 的因子, 即$s = t \times n$, 那么称n 维行向量

$$\langle X, Y\rangle_R := \left(\langle X^1, Y\rangle, \cdots, \langle X^t, Y\rangle\right) \in \mathbb{R}^n. \tag{4.1.17}$$

是X 和Y 的右半张量内积, 这里$X = (X^1, \cdots, X^t)$, $X^i \in \mathbb{R}^n$, $i = 1, \cdots, t$, 并且$\langle\cdot,\cdot\rangle$ 就是普通内积.

情形2: 如果s 是t 的因子, 即$t = s \times n$, 那么称n 维列向量

$$\langle X, Y\rangle_R := \left(\langle Y^{\mathrm{T}}, X^{\mathrm{T}}\rangle_R\right)^{\mathrm{T}} \in \mathbb{R}^n. \tag{4.1.18}$$

是X 和Y 的右半张量内积.

容易证明, 当上面定义中的X 和Y 都是向量时, 有

$$\langle X, Y\rangle_R = X \rtimes Y. \tag{4.1.19}$$

不幸的是, 定义2.3.5 的第2 部分不能推广到右半张量积的情形. 这是由于右半张量积和左半张量积有一个很大的不同: 右半张量积不满足分块乘法规则. 而矩阵的“行乘列”运算方式是由分块乘法规则保证的. 我们用一个简单例子验证右半张量积不满足分块乘法规则.

例 4.1.4 设$A = (a_1, a_2, a_3, a_4)$, $B = (b_1, b_2)^{\mathrm{T}}$, 那么

$$A \rtimes B = (a_1b_1 + a_2b_2\ ,\ a_3b_1 + a_4b_2).$$

如果我们将它们分别等分成$A = (A_1\ \ A_2)$ 和$B = (B_1\ \ B_2)^{\mathrm{T}}$, 根据分块乘法规则, 我们有

$$A_1 \rtimes B_1 + A_2 \rtimes B_2 = (a_1\ \ a_2) \rtimes (b_1) + (a_3\ \ a_4) \rtimes (b_2) = (a_1b_1 + a_3b_2\ \ a_2b_1 + a_4b_2).$$

显然两次运算的结果不一样. □

因此, 尽管左半张量积和右半张量积都是普通矩阵乘法的推广, 在大多数情况下我们还是只使用左半张量积. 由于右半张量积的这个缺点, 在后面的很多应用中左半张量积都不能用右半张量积替代.

下面的命题可以认为是两种半张量积之间的转换规则.

命题 4.1.5 给定$A \in M_{m\times n}$, $B \in M_{p\times q}$. 如果$A \succ_t B$, 那么

$$A \rtimes B = A \ltimes W_{[p,t]} \ltimes B \ltimes W_{[t,q]}. \tag{4.1.20}$$

反之则有

$$A \ltimes B = A \rtimes W_{[t,p]} \rtimes B \rtimes W_{[q,t]}. \tag{4.1.21}$$

如果$A \prec_t B$, 那么

$$A \rtimes B = W_{[m,t]} \ltimes A \ltimes W_{[t,n]} \ltimes B. \tag{4.1.22}$$

反之则有

$$A \ltimes B = W_{[t,m]} \rtimes A \rtimes W_{[n,t]} \rtimes B. \tag{4.1.23}$$

证明 我们只证明(4.1.20). (4.1.21)~(4.1.23) 的证明类似. 利用(3.2.16), 有

$$\begin{aligned} A \rtimes B &= A(I_t \otimes B) = AW_{[p,t]}(B \otimes I_t)W_{[t,q]} \\ &= A \ltimes W_{[p,t]} \ltimes B \ltimes W_{[t,q]}. \end{aligned}$$

□

左半张量内积和右半张量内积之间还有如下的关系.

命题 4.1.6 设X 是一个np 维行向量, Y 是一个p 维列向量, 那么

$$\langle X, Y\rangle_R = \left\langle XW_{[p,n]}, Y\right\rangle_L. \tag{4.1.24}$$

反之则有

$$\langle X, Y\rangle_L = \left\langle XW_{[n,p]}, Y\right\rangle_R. \tag{4.1.25}$$

如果$\dim(X) = p$ 且$\dim(Y) = pn$, 那么

$$\langle X, Y\rangle_R = \left\langle X, W_{[n,p]}Y\right\rangle_L. \tag{4.1.26}$$

反之则有

$$\langle X, Y\rangle_L = \left\langle X, W_{[p,n]}Y\right\rangle_R. \tag{4.1.27}$$

证明 根据命题 4.1.5, (4.1.24)~(4.1.27) 分别是 (4.1.20)~(4.1.23) 的特殊情形.

□

我们可以从多线性映射看出右半张量积的物理意义: 右半张量积也可以自动寻找多维数组的层次关系, 或曰"指针". 设$\sigma : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{s} \to \mathbb{R}^s$ 是一个多重线性映射. 对于$\mathbb{R}^n$ 中给定的一组基底$\{e_1, \cdots, e_n\}$ 及$\mathbb{R}^s$ 中给定的一组基底$\{d_1, \cdots, d_s\}$, 假设

$$\sigma(e_{i_1}, e_{i_2}, \cdots, e_{i_s}) = \sum_{k=1}^{s} \alpha^k_{i_1, \cdots, i_s} d_k.$$

于是结构矩阵M_σ 为

$$M_\sigma = \begin{pmatrix} \alpha^1_{1\cdots 11} & \alpha^1_{1\cdots 12} & \cdots & \alpha^1_{n\cdots nn} \\ \alpha^2_{1\cdots 11} & \alpha^2_{1\cdots 12} & \cdots & \alpha^2_{n\cdots nn} \\ \vdots & \vdots & & \vdots \\ \alpha^n_{1\cdots 11} & \alpha^n_{1\cdots 12} & \cdots & \alpha^n_{n\cdots nn} \end{pmatrix} \tag{4.1.28}$$

对于$X_1, \cdots, X_s \in \mathbb{R}^n$, 将$X_i$ 等同于它的系数向量, 也就是说, 如果$X_i = \sum\limits_{j=1}^{n} x^i_j e_j$, 那么$X_i = (x^i_1, \cdots, x^i_n)^{\mathrm{T}}$. 于是我们有

$$\sigma(X_1, \cdots, X_s) = M_\sigma \ltimes X_1 \ltimes X_2 \ltimes \cdots \ltimes X_s = M_\sigma \rtimes X_k \rtimes X_{k-1} \ltimes \cdots \rtimes X_1. \tag{4.1.29}$$

有时使用右半张量积还是比较方便的. 相应于(3.1.12) 和(3.1.13), 我们有下面的结论.

命题 4.1.7 设$A \in M_{m\times n}$, $X \in M_{n\times q}$, $Y \in M_{p\times m}$, 那么得到乘积的向量形式如下:

$$V_r(YA) = (I_p \otimes A^{\mathrm{T}})V_r(Y) = A^{\mathrm{T}} \rtimes V_r(Y); \tag{4.1.30}$$

$$V_c(AX) = (I_q \otimes A)V_c(X) = A \rtimes V_c(X). \tag{4.1.31}$$

证明 为了证明(4.1.30), 利用表3.2.1,

$$\begin{aligned} V_r(YA) &= V_c(A^{\mathrm{T}}Y^{\mathrm{T}}) \\ &= \left(I_p \otimes A^{\mathrm{T}}\right) V_c(Y^{\mathrm{T}}) \left(I_p \otimes A^{\mathrm{T}}\right) V_r(Y) \\ &= A^{\mathrm{T}} \rtimes V_r(Y). \end{aligned}$$

对于(4.1.31), 利用表3.2.2 及等式(3.3.28), 我们有

$$\begin{aligned} V_c(AX) &= W_{[m,q]}V_r(AX) = W_{[m,q]} \ltimes A \ltimes V_r(X) \\ &= W_{[m,q]} \ltimes A \ltimes W_{[q,p]} \ltimes V_c(X) \\ &= (I_q \otimes A) \ltimes V_c(B) = A \rtimes V_c(B). \end{aligned}$$

□

利用命题4.1.7, 我们还可以证明一些三个矩阵相乘时的公式.

命题 4.1.8

$$V_r(ABC) = (A \otimes C^{\mathrm{T}})V_r(B); \tag{4.1.32}$$

$$V_c(ABC) = (C^{\mathrm{T}} \otimes A)V_c(B). \tag{4.1.33}$$

证明 利用(3.1.12) 和(4.1.29), 有

$$\begin{aligned} V_r(ABC) &= (A \otimes I_q)V_r(BC) = (A \otimes I_q)(I_n \otimes C^{\mathrm{T}})V_r(B) \\ &= (A \otimes C^{\mathrm{T}})V_r(B). \end{aligned}$$

这就证明了(4.1.32).

利用(3.1.13) 和(4.1.30), 有

$$\begin{aligned} V_c(ABC) &= (I_q \otimes A)V_c(BC) = (I_q \otimes A)(C^{\mathrm{T}} \otimes I_n)V_c(B) \\ &= (C^{\mathrm{T}} \otimes A)V_c(B). \end{aligned}$$

这就证明了(4.1.33). □

下面的命题是定义的直接推论.

命题 4.1.9 设$A \in M_{m\times n}$, $B \in M_{n\times m}$, 那么

$$\mathrm{tr}(AB) = V_c^{\mathrm{T}}(A)V_r(B) = V_r^{\mathrm{T}}(B)V_c(A) = V_c^{\mathrm{T}}(B)V_r(A) = V_r^{\mathrm{T}}(A)V_c(B). \tag{4.1.34}$$

利用上面的命题, 以及(4.1.32) 和(4.1.33), 我们有

命题 4.1.10 设A, B, C, D 是使得$ABCD$ 是方阵的矩阵, 那么

$$\mathrm{tr}(ABCD) = V_c^{\mathrm{T}}(A)(B \otimes D^{\mathrm{T}})V_r(C) = V_r^{\mathrm{T}}(C)(B^{\mathrm{T}} \otimes D)V_c(A). \tag{4.1.35}$$

最后, 我们给出两个命题来说明右半张量积的应用. 它们在将一个矩阵多项式表示成它的元素的多项式形式时特别有用.

首先, 引入幂的记号: 设$A \in M_{m\times n}$, 并且$m\%n = 0$ 或$n\%m = 0$, 即A 与自己满足倍维数条件, 那么$A \rtimes A$ 有定义, 进而可以定义A 的右半张量积意义下的高次方如下:

$$A^{\rtimes k} = \underbrace{A \rtimes A \rtimes \cdots \rtimes A}_{k}.$$

注意, 我们约定

$$A^k = \underbrace{A \ltimes A \ltimes \cdots \ltimes A}_{k}.$$

因此, 对右半张量积意义下的高次方,指数上的右半张量积记号是不能省略的. 当然, 当A 为方阵时, 左、右半张量积意义下的乘方都退化为普通矩阵(方阵)的乘方了.

命题 4.1.11 设$Z \in M_n$, 那么

$$Z^k = (F^{\rtimes k}) \ltimes V_c^k(Z), \quad k \geq 1. \tag{4.1.36}$$

其中

$$F = \big(I_n \otimes V_c^{\mathrm{T}}(I_n)\big) \ltimes W_{[n]}.$$

证明 我们归纳证明(4.1.36). 利用(3.3.7), 有

$$Z = \pi_n^c(I_n) \ltimes V_c(Z).$$

令$F = \pi_n^c(I_n)$, 利用(3.3.1) 和(3.3.3) 容易看出, F 具有命题所述的形式. 注意到$F \in M_{n\times n^3}$, 于是$F^{\rtimes k} \in M_{n\times n^{2k+1}}$. 现在假设(4.1.36) 对$k$ 成立. 那么, 利用(3.3.20), 有

$$\begin{aligned}
Z^{k+1} &= Z \ltimes Z^k = F \ltimes V_c(Z) \ltimes F^{\rtimes k} \ltimes V_c^k(Z) \\
&= F \ltimes (I_{n^2} \otimes F^{\rtimes k}) \ltimes V_c(Z) \ltimes V_c^k(Z) \\
&= \big(F(I_{n^2} \otimes F^{\rtimes k})\big) \ltimes V_c^{k+1}(Z) \\
&= F^{\rtimes(k+1)} \ltimes V_c^{k+1}(Z).
\end{aligned}$$

□

命题 4.1.12 设$Z \in M_n$, 那么

$$V_r(Z^k) = \left(E^{\rtimes(k-1)}\right) V_r^k(Z), \quad k \geq 2. \tag{4.1.37}$$

$$V_c(Z^k) = W_{[n]} \left(E^{\rtimes(k-1)}\right) (W_{[n]})^{\otimes k} V_c^k(Z), \quad k \geq 2. \tag{4.1.38}$$

其中

$$E = I_n \otimes V_c^{\mathrm{T}}(I_n).$$

证明 首先证明(4.1.37). 对于$k = 1$, 有

$$V_r(Z^2) = Z \ltimes V_r(Z) = FV_c(Z)V_r(Z) = FW_{[n]}V_r^2(Z) = EV_r^2(Z).$$

现在假设(4.1.37) 对于k 成立, 然后由(3.1.10) 有

$$\begin{aligned} V_r(Z^{k+1}) &= Z \ltimes V_r(Z^k) = E \ltimes V_r(Z) \ltimes E^{\rtimes(k-1)} \ltimes V_r^k(Z) \\ &= \left(E \ltimes (I_{n^2} \otimes E^{\rtimes(k-1)})\right) \ltimes V_r(Z) \ltimes V_r^k(Z) \\ &= \left(E^{\rtimes k}\right) \ltimes V_r^{k+1}(Z). \end{aligned}$$

这样就证明了(4.1.37).

注意到

$$V_r^k(Z) = \underbrace{(W_{[n]} \otimes \cdots \otimes W_{[n]})}_{k} V_c^k(Z) = (W_{[n]})^{\otimes k} V_c^k(Z).$$

用$W_{[n]}$ 左乘(4.1.37) 的两边就得到了(4.1.38). □

4.2 一般矩阵的半张量积

本节我们考虑一般的任意维数矩阵间的左、右半张量积, 简称任意维半张量积. 这是一个非常有趣的话题, 但对它更一般的讨论已经超过了本书的范围, 除非特别说明, 我们只在本节讨论它.

设$a, b \in \mathbb{Z}^+$. 我们记a 和b 的最大公因数为$a \wedge b$, a 和b 最小公倍数为$a \vee b$. 例如, $6 \wedge 8 = 2$, $6 \vee 8 = 24$.

定义 4.2.1 设$A \in M_{m\times n}$, $B \in M_{p\times q}$, 且$\alpha = n \vee p$. 定义A 和B 的左半张量积为

$$A \ltimes B = \left(A \otimes I_{\frac{\alpha}{n}}\right)\left(B \otimes I_{\frac{\alpha}{p}}\right). \tag{4.2.1}$$

定义A 和B 的右半张量积为

$$A \rtimes B = \left(I_{\frac{\alpha}{n}} \otimes A\right)\left(I_{\frac{\alpha}{p}} \otimes B\right). \tag{4.2.2}$$

注意到, 如果$n = p$, 那么一般矩阵的左(右)半张量积就是普通矩阵乘法. 如果$n \vee p = n$ 或者$n \vee p = p$, 即倍维数条件成立, 那么一般矩阵的左(右)半张量积就是前面讨论过的左(右)半张量积. 另外需要说明一下, 只有在本节之外讨论的对矩阵维数有要求的情形才是合法的半张量积. 因此, 这个定义是最具一般性的普适定义.

应当说明的是, 为什么本书不着力研究这种最具一般性的矩阵乘法? 其主要原因是, 到目前为止, 我们尚未找到它的任何重要应用. 我们对满足倍维数条件的两个矩阵的半张量积的兴趣主要不是它的理论结果, 虽然我们得到许多简洁有趣的结果, 我们的兴趣主要来源于它明确的物理背景和广泛的实用性.

命题 4.2.2 半张量积满足

(1) 分配律

$$(A+B) \ltimes C = (A \ltimes C) + (B \ltimes C), \tag{4.2.3}$$

$$(A+B) \rtimes C = (A \rtimes C) + (B \rtimes C), \tag{4.2.4}$$

$$C \ltimes (A+B) = (C \ltimes A) + (C \ltimes B), \tag{4.2.5}$$

$$C \rtimes (A+B) = (C \rtimes A) + (C \rtimes B). \tag{4.2.6}$$

(2) 结合律

$$(A \ltimes B) \ltimes C = A \ltimes (B \ltimes C), \tag{4.2.7}$$

$$(A \rtimes B) \rtimes C = A \rtimes (B \rtimes C). \tag{4.2.8}$$

证明 分配律显然成立. 我们只证明左半张量积的结合律.

设$A \in M_{m\times n}$, $B \in M_{p\times q}$, $C \in M_{s\times t}$. 我们首先要求(4.2.7)~(4.2.8) 的两边在形式上都可以表示成

$$(A \otimes I_\alpha)(A \otimes I_\beta)(A \otimes I_\gamma). \tag{4.2.9}$$

这可根据下面的等式:

$$(A \otimes I_m)(B \otimes I_n) \otimes I_s = (A \otimes I_{m+s})(B \otimes I_{n+s}),$$

直接计算得到.

从乘法的定义可以看出, 不管按哪种顺序作乘法, 最后它们都要找到最小的自然数α, β, γ, 使得

$$\begin{cases} \alpha n = \beta p, \\ \beta q = \gamma s \end{cases} \tag{4.2.10}$$

成立. 如果我们能够说明(4.2.10) 的最小解是唯一的, 那也就证明了结论. 既然$\alpha nq = \gamma sp$, 可设$nq \wedge sp = h$, 那么

$$\alpha = \mu\frac{sp}{h}, \quad \gamma = \mu\frac{nq}{h}.$$

于是有

$$\beta = \frac{\mu sn}{h} = \frac{sn}{h/\mu}.$$

现在最小的μ 使得$c = (h/\mu)$ 是nq, sp 和sn 的最大公倍数. 令

$$c = nq \wedge sp \wedge sn,$$

那么就有(4.2.7) 及(4.2.8) 的两边都等于(4.2.4), 并且

$$\alpha = \frac{sp}{c}, \quad \beta = \frac{sn}{c}, \quad \gamma = \frac{nq}{c}.$$

□

几乎所有普通矩阵乘法的性质对左(右) 半张量积仍然成立. 例如, 我们有

命题 4.2.3 (1)

$$\begin{cases}(A \ltimes B)^{\mathrm{T}} = B^{\mathrm{T}} \ltimes A^{\mathrm{T}},\\ (A \rtimes B)^{\mathrm{T}} = B^{\mathrm{T}} \rtimes A^{\mathrm{T}}.\end{cases} \tag{4.2.11}$$

(2) 如果$M \in M_{m\times pn}$, 那么

$$\begin{cases}M \ltimes I_n = M,\\ M \rtimes I_n = M.\end{cases} \tag{4.2.12}$$

如果$M \in M_{pm\times n}$, 那么

$$\begin{cases}I_m \ltimes M = M,\\ I_m \rtimes M = M.\end{cases} \tag{4.2.13}$$

在下面条目中, A, B 是两个方阵.

(3) $A \ltimes B$ 和$B \ltimes A$ ($A \rtimes B$ 和$B \rtimes A$) 具有相同的特征函数.

(4)

$$\begin{cases}\mathrm{tr}(A \ltimes B) = \mathrm{tr}(B \ltimes A),\\ \mathrm{tr}(A \rtimes B) = \mathrm{tr}(B \rtimes A).\end{cases} \tag{4.2.14}$$

(5) 如果A 和B 都是正交阵(上三角阵, 下三角阵或对角阵), 那么$A\ltimes B$ ($A\rtimes B$) 也是.

(6) 如果A 或B 可逆, 那么$A\ltimes B\sim B\ltimes A$ ($A\rtimes B\sim B\rtimes A$).

(7) 如果A 和B 都可逆, 那么

$$\begin{cases}(A\ltimes B)^{-1}=B^{-1}\ltimes A^{-1},\\(A\rtimes B)^{-1}=B^{-1}\rtimes A^{-1}.\end{cases}\tag{4.2.15}$$

(8) 乘积的行列式为

$$\begin{cases}\det(A\ltimes B)=[\det(A)]^{\frac{\alpha}{n}}[\det(B)]^{\frac{\alpha}{p}},\\\det(A\rtimes B)=[\det(A)]^{\frac{\alpha}{n}}[\det(B)]^{\frac{\alpha}{p}}.\end{cases}\tag{4.2.16}$$

推论 4.2.4 设$A\in M_{m\times n}$, $B\in M_{p\times q}$, 那么

$$C=A\ltimes B=(C^{ij}),\quad i=1,\cdots,m;\ j=1,\cdots,q.\tag{4.2.17}$$

其中

$$C^{ij}=A^i\ltimes B_j.$$

这里A^i 表示A 的第i 行, B_j 表示B 的第j 列.

注意到(4.2.17) 就是矩阵乘法形式. 唯一的不同就是, 前一个矩阵的行和后一个矩阵的列之间的"内积" 是一个块, 而不是一个数.

注 从本节开始, 对于任意两个矩阵$A_{m\times n}$ 和$B_{p\times q}$, 我们将只考虑

$$AB=A\ltimes B.$$

当$n=p$ 时, 它就是普通矩阵乘法; 当$n\vee p=\max\{n,p\}$ 时, 它就是在第2, 3章中讨论的左半张量积. 因此对于左半张量积, 我们可以省略$\ltimes$. 但是在少数情况下, 我们仍然使用$\ltimes$ 这个记号. 其可能的原因是:强调它的半张量积特性, 或避免读者可能的概念混淆. 实际上, 读者应该让自己确信, 普通矩阵乘法完全可以由左半张量积取代.

设$A\in M_{m\times n}$ 我们定义A 的左半张量幂为

$$\begin{cases}A^1=A,\\A^{k+1}=A\ltimes A^k,\quad k\geq 1.\end{cases}$$

同样可以定义A 的右半张量幂为

$$\begin{cases} A^{\rtimes 1} = A, \\ A^{\rtimes (k+1)} = A \rtimes A^{\rtimes k}, \quad k \geq 1. \end{cases}$$

至于A^k ($A^{\rtimes k}$) 的维数, 设$m \wedge n = t$. 令$m = m_0 t$, $n = n_0 t$, 那么m_0 和n_0 互质. 很容易通过数学归纳证明$A^k \in M_{m_0^k t \times n_0^k t}$ ($A^{\rtimes k} \in M_{m_0^k t \times n_0^k t}$).

4.3　半张量代数

本节中讨论的左(右)半张量积仍在满足倍维数条件下定义.

本节讨论一组方阵在左半张量积下构成的代数结构, 称为半张量代数. 它有许多深刻的性质, 这里给出的只是一些初步的讨论. 另外, 在本节的讨论中, 左半张量积与右半张量积没有区别, 因此将本节的左半张量积用右半张量积代替, 全部结论均成立.

定义一组方阵如下:

$$W_n = \{A \in M_{n^k \times n^k} \mid k = 0, 1, 2, \cdots\}.$$

注意到W_n 中的任何两个矩阵均满足倍维数条件, 因此左半张量积是定义好的. 但是W_n 中的两个矩阵一般具有不同维数, 因此无法定义加法. 于是我们需要如下定义:

定义 4.3.1　设$A, B \in W_n$, 称A, B 是等价的, 记作$A \sim B$, 如果存在一个单位阵I_{n^s} 使得

$$A \otimes I_{n^s} = B \quad 或 \quad B \otimes I_{n^s} = A.$$

显然, $\sim$ 是一个等价条件, 因此我们可以对等价作分割, 形成等价类集合. 记

$$\mathcal{W}_n = W_n / \sim.$$

设$A \in W_n$, 记A 的等价类为$[A]$, 则$[A] \in \mathcal{W}_n$.

由W_n 的结构可知, 设$A, B \in W_n$, 则一定存在t (t 可表示为$t = n^s$), 使得$A \prec_t B$ 或$A \succ_t B$. 下面我们在$\mathcal{W}_n$ 上定义加、减及乘法运算.

定义 4.3.2　设$A, B \in W_n$, 并且

(1) 设$A \prec_t B$, 则定义加法(“$+$”) 及减法(“$-$”) 分别如下:

$$[A] + [B] = [A \otimes I_t + B]. \tag{4.3.1}$$

$$[A] - [B] = [A \otimes I_t - B]. \tag{4.3.2}$$

(2) 设$A \succ_t B$, 则定义加法及减法分别如下:

$$[A] + [B] = [A + B \otimes I_t]. \tag{4.3.3}$$

$$[A] - [B] = [A - B \otimes I_t]. \tag{4.3.4}$$

(3) 定义乘法如下

$$[A] \ltimes [B] = [A \ltimes B]. \tag{4.3.5}$$

当然, 我们必须检验上述定义的加、减、乘法是定义好的, 即它不依赖于代表元的选择. 一旦检验了这一点, 下面的基本性质就是显见的了. 我们将这些检验工作留给读者(参见习题4.6, 4.7).

命题 4.3.3 (1)

$$r[A] = [rA], \quad r \in \mathbb{R}. \tag{4.3.6}$$

(2) 分配律

$$([A] \pm [B]) \ltimes [C] = [A \ltimes C] \pm [B \ltimes C]; \tag{4.3.7}$$

$$[C] \ltimes ([A] \pm [B]) = [C \ltimes A] \pm [C \ltimes B]. \tag{4.3.8}$$

(3) 结合律

$$([A] \ltimes [B]) \ltimes [C] = [A] \ltimes ([B] \ltimes [C]). \tag{4.3.9}$$

下面考虑W_n 的一组子集如下:

$$W_n^m = \{ A \in M_{n^k \times n^k} \mid 0 \le k \le m \}.$$

相应地, 我们定义$\mathcal{W}_n$ 的一组子集如下:

$$\mathcal{W}_n^m = \{ [A] \mid A \in W_n^m \}.$$

下面的几个命题都是比较简单的, 我们均将证明留给读者.

命题 4.3.4 $(\mathcal{W}_n, +)$ 是一个Abel 群, 且$\mathcal{W}_n^m < \mathcal{W}_n$, 即$\mathcal{W}_n^m$ 是$\mathcal{W}_n$ 的一个子群.

命题 4.3.5 $(\mathcal{W}_n, +, \ltimes)$ 是一个非交换的结合代数, 有单位元1. $\mathcal{W}_n^m$ 是它的一个子代数.

由于$(\mathcal{W}_n,+)$ 是$\mathbb{R}$ 上的一个向量空间, 对$[A],[B]\in\mathcal{W}_n$ 我们可以定义如下的李括号运算:

$$[[A],[B]]=[A]\ltimes[B]-[B]\ltimes[A]. \tag{4.3.10}$$

于是有如下结论:

命题 4.3.6　对由(4.3.10) 定义的李括号$\mathcal{W}_n$ 成为一个李代数. $\mathcal{W}_n^m$ 是它的一个李子代数.

如果A 可逆, 则可定义$[A]^{-1}\in\mathcal{W}_n$, 它满足

$$[A]\ltimes[A]^{-1}=[1]. \tag{4.3.11}$$

命题 4.3.7　如果A 可逆, 则$[A]^{-1}\in\mathcal{W}_n$ 存在且唯一, 它满足

$$[A]^{-1}=[A^{-1}]. \tag{4.3.12}$$

定义$\mathcal{W}_n$ 的一组子集如下:

$$\mathcal{G}_n=\left\{[A]\in\mathcal{W}_n \mid \text{如果 } [A]^{-1} \text{ 存在}\right\}.$$

我们可以定义它相应的一组子集如下:

$$\mathcal{G}_n^m=\mathcal{W}_n^m\cap\mathcal{G}_n.$$

于是我们有

命题 4.3.8　$(\mathcal{G}_n,\ltimes)$ 是一个群, $\mathcal{G}_n^m$ 是$\mathcal{G}_n$ 的一个子群.

粗略地说, $\mathcal{G}_n$ 对应于一个(不定维数的)一般线性群$GL(\cdot,\mathbb{R})$. 那么, 我们也能定义它的相应于特殊线性群$SL(\cdot,\mathbb{R})$, 正交群$O(\cdot,\mathbb{R})$ 等的子群. 我们把这些留给读者(见习题4.8).

设$[A]\in\mathcal{W}_n$ 我们可以定义它的指数映射

$$\exp([A]):=[e^A]. \tag{4.3.13}$$

命题 4.3.9　(4.3.13) 中定义的$\exp([A])$ 是定义好的, 并且$\exp([A])\in\mathcal{G}_n$.

证明　后半部分是显然的, 我们只证明前半部分. 设$B=A\otimes I$, 我们只要证明$[e^B]=[e^A]$ 即可.

$$\begin{aligned} e^B &= e^{A\otimes I}=I+(A\otimes I)+\tfrac{1}{2}(A\otimes I)^2+\cdots \\ &= I+(A\otimes I)+\tfrac{1}{2}(A^2\otimes I)+\cdots \\ &= (I+A+\tfrac{1}{2}A^2+\cdots)\otimes I=e^A\otimes I. \end{aligned}$$

于是$[e^B]=[e^A]$. □

最后, 我们讨论$\mathcal{W}_n$ 上的拓扑结构. 对$[A]\in\mathcal{W}_n$, 我们定义一个范数如下:

$$\|[A]\|=\frac{1}{\sqrt{k}}\|V_c(A)\|, \tag{4.3.14}$$

这里$A\in M_{k\times k}$, $\|V_c(A)\|$ 为$\mathbb{R}^{k^2}$ 上的欧氏范数, 即

$$\|V_c(A)\|=\sqrt{\sum_{i=1}^{k}\sum_{j=1}^{k}(a_{ij})^2}.$$

命题 4.3.10 (4.3.14) 中定义的$\|[A]\|$ 是$\mathcal{W}_n$ 上定义好的一个范数.

证明 首先证明它是定义好的, 设$A\in M_{k\times k}$ 且$B=A\otimes I_s$, 当分别用A, B 作代表元时, 分别有

$$\|[A]\|=\frac{1}{\sqrt{k}}\sqrt{\sum_{i=1}^{k}\sum_{j=1}^{k}(a_{ij})^2}.$$

$$\|[B]\|=\frac{1}{\sqrt{ks}}\sqrt{s\sum_{i=1}^{k}\sum_{j=1}^{k}(a_{ij})^2}=\|[A]\|.$$

所以它与代表元无关, 即它是定义好的.

要证明它是范数, 只要将范数定义条件中有关的(有限个)元素的代表元选在同一维数的矩阵空间里, 则可看出它满足范数的要求. □

由命题 4.3.10 可知, 在所选范数下 $\mathcal{W}_n$ 是一个赋范空间. 但是, 它不是一个 Banach 空间. 这可由下面的例子看出:

例 4.3.11 构造一个序列$[A]_i=[A_i]$ 如下:

$$A_1=\begin{bmatrix}1&0\\0&-1\end{bmatrix},\quad A_2=\begin{bmatrix}1+\frac{1}{2}&0&0&0\\0&1-\frac{1}{2}&0&0\\0&0&-1+\frac{1}{2}&0\\0&0&0&-1-\frac{1}{2}\end{bmatrix},$$

$$\begin{aligned}A_3 &= \mathrm{diag}(1+\tfrac{1}{2}+\tfrac{1}{4},1+\tfrac{1}{2}-\tfrac{1}{4},1-\tfrac{1}{2}+\tfrac{1}{4},1-\tfrac{1}{2}-\tfrac{1}{4},\\ &\qquad -1+\tfrac{1}{2}+\tfrac{1}{4},-1+\tfrac{1}{2}-\tfrac{1}{4},-1-\tfrac{1}{2}+\tfrac{1}{4},-1-\tfrac{1}{2}-\tfrac{1}{4}),\\ &\qquad \cdots\end{aligned}$$

容易看出, $\{[A_i]\}$ 是一个Cauchy 序列, 因为

$$\|[A_m]-[A_n]\| \le \frac{1}{2^{m-1}}, \quad n \ge m.$$

现在设该Cauchy 序列收敛于$[A_0]$, 其中A_0 为$2^m \times 2^m$ 矩阵. 由于A_i 中元素的对称分布, 容易证明

$$\|[A_n]-[A_0]\| \ge \|[A_n]-[A_m]\|, \quad n > m.$$

因此, 只要证明

$$\lim_{n\to\infty} \|[A_n]-[A_m]\| \ne 0.$$

即得矛盾. 不难证明

$$\|[A_m]-[A_n]\| \ge \frac{1}{2^{m+1}}, \quad n \ge m.$$

故得矛盾. □

4.4 注释与参考

矩阵的张量积虽然对任何两个矩阵均有定义, 但它与普通乘法是风马牛不相及的两种运算. 矩阵的半张量积则是对普通乘法的推广, 它的理论意义是显见的. 进一步的研究也是有价值的. 本书的一个重要议题是半张量积的应用. 目前, 得到广泛应用的是倍维数的情况. 因此, 我们集中研究倍维数的情况.

左半张量积可以用类似普通乘法的行列相乘来定义, 因此, 它比右半张量积更加自然. 同时,它满足分块乘法规则,这对应用十分有利. 因此,本书主要研究左半张量积. 除本章或其他章节个别另加说明的地方外, 本书其他章节均只涉及倍维数的左半张量积.

代数是抽象代数中的一个重要概念, 直观地说, 它是向量空间加上满足双线性性的向量积. 前面讨论过的李代数, 就是一种特殊的代数结构. 有关概念可参考一些标准代数教程, 例如文献[67], [77], 有关李代数基础也可参考文献[99].

关于赋范空间及Banach 空间等概念可参考一些标准泛函分析教程, 例如文献[39].

习 题 四

4.1 证明(4.1.6)和(4.1.7).

4.2 证明(4.1.8).

4.3 证明(4.1.10)~(4.1.12).

4.4 在立方阵乘法中, 设

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_m \end{bmatrix},$$

其中X_i, $i = 1, \cdots, n$ 是$n \times n$ 矩阵. 设$b \in \mathbb{R}^n$, X 的二次型的定义如文献[107] 中所示为

$$b^{\mathrm{T}} X b := \begin{bmatrix} b^{\mathrm{T}} X_1 b \\ \vdots \\ b^{\mathrm{T}} X_m b \end{bmatrix}.$$

证明它可以表示成

$$b^{\mathrm{T}} X b = b^{\mathrm{T}} \rtimes (X b) = (b^{\mathrm{T}} \rtimes X) b.$$

4.5 给出$V_r(AX)$ 及在右半张量积下的表达式. (参见命题4.1.7 及习题3.3)

4.6 证明定义4.3.2 中所定义的等价类加,减,乘法是定义好的.

4.7 证明命题4.3.3, 4.3.4, 4.3.5, 4.3.6, 4.3.7, 4.3.8.

4.8 (a) 定义$\mathcal{O}_n = \{[A] \in \mathcal{G}_n \mid A^{\mathrm{T}} = A^{-1}\}$. 证明上述集合是定义好的, 且它是$\mathcal{G}_n$ 的一个子群.

(b) 定义$\mathcal{S}_n = \{[A] \in \mathcal{G}_n \mid \det(A) = 1\}$. 证明上述集合是定义好的, 且它是$\mathcal{G}_n$ 的一个子群.

第 5 章　多项式运算的半张量积方法

本章首先介绍由多项式构成的多维映射的半张量积表示. 接着, 利用半张量积方法引入在张量表示下的微分形式及其计算公式. 然后, 我们考虑多项式的两种不同的基底: 自然基底和张量积形式的冗余基底, 并且给出它们之间的变换公式. 继之, 作为一个应用实例, 将这些结果应用于逆映射的Taylor 展开.

在本章第5 节给出一般多元多项式的微分公式, 它在以后的许多应用中起着关键作用. 最后一节给出李导数的计算公式.

多项式的半张量积表示为我们提供了一种解决非线性问题的线性方法, 并且相关运算很容易通过计算机来实现.

5.1　多项式的半张量积表示

设$x=(x_1,\cdots,x_n)$ 是$\mathbb{R}^n$ 中的一个坐标. 记变元为x 的k 次齐次多项式全体的集合为B_n^k. 为方便起见, 令$B_n^0=\mathbb{R}$, 它代表“0” 次多项式, 即常数项全体.

在前面曾提到过x^k 是B_n^k 的一组生成基, 也就是说, 它是包含B_n^k 的一组基底的集合, 因此$f(x)\in B_n^k$ 可以表示成

$$f(x)=F\ltimes x^k,\quad F\in M_{1\times n^k}. \tag{5.1.1}$$

但是这组生成基是冗余的. 也就是说, 这组生成基里的元素并不是线性无关的. 因此(5.1.1) 中的F 不唯一.

例 5.1.1　设$f(x)\in B_2^3$, 且

$$f(x)=x_1^3+x_1^2x_2-2x_1x_2^2-x_2^3. \tag{5.1.2}$$

注意到

$$x=(x_1,\ x_2)^{\mathrm{T}},\quad x^2=(x_1^2,\ x_1x_2,\ x_2x_1,\ x_2^2)^{\mathrm{T}},$$

$$x^3=(x_3,\ x_1^2x_2,\ x_1x_2x_1,\ x_1x_2^2,\ x_2x_1^2,\ x_2x_1x_2,\ x_2^2x_1,\ x_2^3)^{\mathrm{T}}.$$

因此$f(x)$ 可以表示成

$$f(x)=(1\ 1\ 0\ \ -2\ 0\ 0\ 0\ \ -1)\ltimes x^3. \tag{5.1.3}$$

由于生成基是冗余的, 所以表示不唯一. 例如, $f(x)$ 的另一个表示形式为

$$f(x)=(1\ 1\ 0\ \ -1\ 0\ \ -1\ 0\ \ -1)\ltimes x^3. \tag{5.1.4}$$

□

上面多项式的半张量积表示形式的优点在于半张量积具有结合律，在应用中会更加自然和方便. 例如，我们可以用类似于一元多项式因式分解的方法来处理多元多项式，如

$$F_3 \ltimes x^3 + F_5 \ltimes x^5 = (F_3 + F_5 \ltimes x^2) \ltimes x^3.$$

既然x^k 是冗余的，它使得F 的表示不唯一，因此我们想寻找一种方法可以唯一地确定F. 回顾在线性代数中$f(x)$ 的二次型表示为

$$f(x) = x^{\mathrm{T}} F x,$$

这里F 不唯一. 但是如果要求F 对称，那它就是唯一的. 同样，我们可以定义(5.1.1) 的一个对称表示.

定义 5.1.2 设$f(x) \in B_n^k$，并且F 是(5.1.1) 中$f(x)$ 的系数. 称F 是一组对称系数，如果多指标$(i_1, \cdots, i_k)$ 按照索引$Id(i_1, \cdots, i_k; n, \cdots, n)$ 排列有

$$F_{i_{\sigma(1)}, \cdots, i_{\sigma(k)}} = F_{i_1, \cdots, i_k}, \quad \forall \sigma \in \mathbf{S}_k,$$

其中$\mathbf{S}_k$ 表示k 元置换群.

注 (1) 如果(5.1.1) 中的F 对称，那么它是唯一的.

(2) 如果$f(x)$ 是一个二次型，并且F 是一组对称系数，即$x \in \mathbb{R}^n$ 且

$$f(x) = Fx^2 = (F_{11}, F_{12}, \cdots, F_{1n}, \cdots, F_{n1}, F_{n2}, \cdots, F_{nn})x^2,$$

其中$F_{ij} = F_{ji}$，那么$V_c^{-1}(F) = V_r^{-1}(F)$ 是一个对称矩阵. 因此，对称系数组是对称矩阵的自然推广.

(3) 实际上，我们可以将$f(x)$ 看作一个张量$f \in T^k(V)$，即

$$f(z_1, \cdots, z_k) := F \ltimes z_1 \ltimes \cdots \ltimes z_k$$

在$z_1 = \cdots = z_k = x$ 时的特殊情形. 因此齐次多项式具有张量结构. 而用半张量积方法进行张量计算的优越性在第2 章中已有讨论.

在后面的讨论中，置换群是一个基本工具，我们在第1 章中也已介绍过，这里再作一点简单回顾. 置换群$\mathbf{S}_n$ 是由n 个元素的全体置换组成. 因此S_n 共有$n!$ 个置换. 例如，$\mathbf{S}_5$ 有$5! = 120$ 个置换. 下面给出一个例子说明置换和指标置换.

例 5.1.3 (1) 设$\sigma_1, \sigma_2 \in \mathbf{S}_6$ 为

$$\sigma_1 = (123)(456), \quad \sigma_2 = (15642).$$

那么, 其乘积为

$$\sigma_2\sigma_1 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 2 & 3 & 1 & 5 & 6 & 4 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ 1 & 3 & 5 & 6 & 4 & 2 \end{bmatrix}.$$

简记为: $\sigma_2\sigma_1 = (23546)$.

(2) 设$F_{i_1,i_2,i_3,i_4,i_5,i_6}$ 由索引$Id(i_1,i_2,i_3,i_4,i_5,i_6;n,n,n,n,n,n)$ 标注. 例如, $F_{i_1,i_2,i_3,i_4,i_5,i_6} = F_{1,3,2,3,1,1}$, 即$(i_1,i_2,i_3,i_4,i_5,i_6) = (1,3,2,3,1,1)$,

那么用F_σ 代表下标经换置后的元素$F_{i_{\sigma(1)},\cdots,i_{\sigma(k)}}$. 以$a = F_{1,3,2,3,1,1}$ 为例, 当$\sigma = \sigma_1$ 时, $\sigma_1(1) = 2$, 而$i_2 = 3$, 故a 的相应元素第一个指标为"3", $\sigma_1(2) = 3$, 而$i_3 = 2$, 故该元素第二个指标为"2", 如此等等, 最后可得$F_{\sigma(1)} = F_{3,2,1,1,1,3}$. 同样, 可以检验$F_{\sigma(2)} = F_{1,1,2,3,1,3}$, $F_{\sigma(2)\sigma_1} = F_{1,2,1,1,1,3}$. □

那么对于(5.1.1) 中给出的$f(x)$ 的表示, 如何将它转化为对称表示呢? 一个简单的方法就是合并同类项, 但是一般需要一个公式机械地完成. 如果有这样的一个公式, 那么就可以很方便地进行理论研究和计算机实现. 为此, 我们需要引入一些新的概念和记号.

定义 5.1.4　设$I = (I_1, I_2, \cdots, I_k) \in Id(i_1, i_2, \cdots, i_k; \underbrace{n, \cdots, n}_{k})$, 那么

(1) 定义I 的置换集合P_I 为

$$\begin{aligned} P_I \quad = \quad & \{J \in Id(i_1, i_2, \cdots, i_k; \underbrace{n, \cdots, n}_{k}) \mid \\ & (J_1, \cdots, J_k) = (I_{\sigma(1)}, \cdots, I_{\sigma(k)}),\ \sigma \in \mathbf{S}_k\}. \end{aligned}$$

(2) 定义I 中各个指标出现的频率为$C_I = (c_1, \cdots, c_n) \in \mathbb{Z}_+^n$, 其中$c_j$ 是j 在$\{I_1, I_2, \cdots, I_k\}$ 中出现的次数.

记P_I 的大小(即P_I 的元素个数)为$|P_I|$, 并且记

$$C_I! = \prod_{j=1}^{n} c_j!.$$

那么显然有

$$|P_I| = \frac{k!}{C_I!}. \tag{5.1.5}$$

例 5.1.5 设$k=4$, $n=5$, $I=(2,5,2,3)\in Id(i_1,i_2,i_3,i_4;n^4)$, 那么

$$P_I=\left\{\begin{array}{cccccc}(2,2,3,5) & (2,2,5,3) & (2,3,2,5) & (2,3,5,2) & (2,5,2,3) & (2,5,3,2)\\ (3,2,2,5) & (3,2,5,2) & (3,5,2,2) & (5,2,2,3) & (5,2,3,2) & (5,3,2,2)\end{array}\right\},$$

由于I 中没有1, 因此$c_1=0$. 同理有, $c_2=2$, $c_3=1$, $c_4=0$, $c_5=1$. 因此我们有$C_I=(0,2,1,0,1)$, 于是

$$|P_I|=\frac{k!}{C_I!}=\frac{4!}{0!2!1!0!1!}=12.$$

□

下面的命题是上面的定义和记号的直接推论.

命题 5.1.6 设$P(x)=Fx^k$, $x\in\mathbb{R}^n$ 是一个k 次齐次多项式, $P(x)=\tilde{F}x^k$ 是它的对称表示. 设F 和$\tilde{F}$ 都是按照索引$Id(i_1,i_2,\cdots,i_k;\underbrace{n,\cdots,n}_{k})$ 排列的, 那么对于$I=(i_1,\cdots,i_k)\in Id(i_1,i_2,\cdots,i_k;\underbrace{n,\cdots,n}_{k})$ 有

$$\tilde{F}_I=\frac{1}{|P_I|}\sum_{J\in P_I}F_J. \tag{5.1.6}$$

定义 5.1.7 设 $F=\{F_{i_1,\cdots,i_k}\}\in\mathbb{R}^{n^k}$ 是由 n^k 个数按照索引 $Id(i_1,\cdots,i_k;\underbrace{n,\cdots,n}_{k})$ 排列的集合. 我们称由 (5.1.6) 定义的映射 $\tilde{F}=\psi_n^k(F):\mathbb{R}^{n^k}\to\mathbb{R}^{n^k}$ 为 $\mathbb{R}^{n^k}$ 上的对称化算子.

根据定义显然有ψ_n^1 及ψ_n^0 是恒等映射.

下面的命题是定义的一个直接推论.

命题 5.1.8 一个k 次多项式

$$P(x)=P_0+P_1x+P_2x^2+\cdots+P_kx^k,\quad x\in\mathbb{R}^n$$

恒等于零, 当且仅当

$$\psi_n^i(P_i)=0,\quad i=0,1,\cdots,k.$$

设 $s\geq 2$, 定义一个 $n^s\times n^s$ 矩阵 Ψ_n^s 如下: 用指标 $\{i_1,\cdots,i_s\}$ 依索引 $Id(i_1,\cdots,i_s;n,\cdots,n)$ 标注它的行和列, 设 $J=(J_1,\cdots,J_s)$ 为其行指标, $I=$

$(I_1, \cdots, I_s)$ 为其列指标. 记 $\Psi_n^s = (\psi_{J,I})$, 且定义其元素为

$$\psi_{J,I} = \begin{cases} 0, & J \notin P_I, \\ \frac{1}{|P_I|}, & J \in P_I. \end{cases}$$

定义$\Psi_n^0 = 1$, $\Psi_n^1 = I_n$, 那么, 这个Ψ_n^k 可实现映射ψ_n^k:

命题 5.1.9　一个k 次多项式

$$P(x) = P_0 + P_1 x + P_2 x^2 + \cdots + P_k x^k, \quad x \in \mathbb{R}^n,$$

其对称系数表示为

$$P(x) = \tilde{P}_0 + \tilde{P}_1 x + \tilde{P}_2 x^2 + \cdots + \tilde{P}_k x^k,$$

这里

$$\tilde{P}_i = P_i \Psi_n^i, \quad i = 0, 1, \cdots, k. \tag{5.1.7}$$

$P(x)$ 恒等于零, 当且仅当$P_i \Psi_n^i = 0$, $i = 0, 1, \cdots, k$.

例 5.1.10　(1) 设$n = 3$, $k = 2$, 则Ψ_3^2 为

$$\Psi_3^2 = \begin{array}{c} \begin{matrix} 11 & 12 & 13 & 21 & 22 & 23 & 31 & 32 & 33 \end{matrix} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} \\ (11) \\ (12) \\ (13) \\ (21) \\ (22) \\ (23) \\ (31) \\ (32) \\ (33) \end{array}$$

(2) 设$n=2$, $k=3$, 则Ψ_2^3 为

$$
\begin{array}{c}
\begin{array}{cccccccc}
111 & 112 & 121 & 122 & 211 & 212 & 221 & 222
\end{array}
\\
\Psi_3^2 = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\
0 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\
0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\
0 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\
0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\
0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{array}{c}
(111)\\(112)\\(121)\\(122)\\(211)\\(212)\\(221)\\(222)
\end{array}
\end{array}
$$

(3) 再次考虑多项式(5.1.2), 易看出(5.1.3) 和(5.1.4) 都不对称. 任选这两个表达式中的一个, 取其系数F, 利用公式(5.1.7) 可得(5.1.2) 的对称系数表示是

$$f(x)=\begin{bmatrix}1 & \frac{1}{3} & \frac{1}{3} & -\frac{2}{3} & \frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} & -1\end{bmatrix}\ltimes x^3.$$

□

设$P(x)$, $x\in\mathbb{R}^n$ 是实系数多项式在$\mathbb{R}$ 上构成的向量空间, 那么它是所有不同次数的齐次多项式空间的直和, 可以表示成

$$P=B_n^0\oplus B_n^1\oplus B_n^2\oplus\cdots. \tag{5.1.8}$$

命题 5.1.11 设$P(x)\in B_n^p$, $Q(x)\in B_n^q$, 并且它们相应的系数分别是M_P和M_Q, 那么它们的乘积$P(x)Q(x)\in B_n^{p+q}$ 的系数是

$$M_{PQ}=M_P\ltimes M_Q. \tag{5.1.9}$$

证明

$$
\begin{aligned}
M_P\ltimes M_Q\ltimes x^{p+q} &= M_P\ltimes M_Q\ltimes x^q\ltimes x^p=M_P\ltimes(M_Q\ltimes x^q)\ltimes x^p\\
&=(M_Q\ltimes x^q)\ltimes(M_P\ltimes x^p)=M_{PQ}\ltimes x^{p+q}.
\end{aligned}
$$

□

下面的命题是上述结果的一个直接推论.

命题 5.1.12 设$M_P\in\mathbb{R}^{n^p}$ 和$M_Q\in\mathbb{R}^{n^q}$ 是两个行向量, 那么

$$M_P\ltimes M_Q\ltimes x^{p+q}=M_Q\ltimes M_P\ltimes x^{p+q}. \tag{5.1.10}$$

由于x^k 是B_n^k 的一组冗余生成基, 故由(5.1.10) 不能得到$M_P \ltimes M_Q = M_Q \ltimes M_P$.

根据命题5.1.11, 所有基本的代数公式都适用于多元多项式. 例如, 对于公式$(a+b)(a-b)=a^2-b^2$, 同样也有

$$(F_1 \ltimes x + F_2 \ltimes x^2)(F_1 \ltimes x - F_2 \ltimes x^2) = F_1^2 \ltimes x^2 - F_2^2 \ltimes x^4$$

等.

对于(5.1.10), 我们有下面的推广:

命题 5.1.13　设$F \in R^{n^\alpha}$ 是一个行向量, 且$t \geq \alpha$, 那么

$$(G \ltimes F) \ltimes x^t = F \ltimes x^\alpha \ltimes G \ltimes x^{t-\alpha}. \tag{5.1.11}$$

实际上, (5.1.11) 中的G 可以是任意矩阵.

为了体现这种表示的优势, 我们考虑下面的例子:

例 5.1.14　设$x \in \mathbb{R}^n$, 并给定多项式

$$P(x) = A_0 + A_1x + A_2x^2 + \cdots + A_tx^t.$$

我们想知道它是否有一个线性因子$Q(x) = C - Dx$, 其中$C \neq 0$.

不失一般性, 我们可设$C = 1$. 假设我们已经有

$$\begin{aligned}&(1 - Dx)(B_0 + B_1x + \cdots + B_{t-1}x^{t-1})\\&= A_0 + A_1x + A_2x^2 + \cdots + A_tx^t.\end{aligned} \tag{5.1.12}$$

比较(5.1.12) 两边的系数可以得到

$$\begin{cases}B_0 = A_0,\\B_k = A_k + DB_{k-1}, \quad k = 1, \cdots, t-1.\end{cases} \tag{5.1.13}$$

因此要使(5.1.12) 成立, 当且仅当由上述递推式定义的B_{t-1} 满足

$$-DB_{t-1}x^t = A_tx^t. \tag{5.1.14}$$

将由(5.1.13) 算出的B_{t-1} 代入(5.1.14), 得

$$(A_t + DA_{t-1} + D^2A_{t-2} + \cdots + D^tA_0)x^t = 0. \tag{5.1.15}$$

□

利用上例及命题5.1.9, 我们得到下面的命题:

命题 5.1.15 多项式$P(x) = A_0 + A_1x + A_2x^2 + \cdots + A_tx^t$ 有一个线性因子$1 - Dx$, 当且仅当

$$(A_t + DA_{t-1} + D^2A_{t-2} + \cdots + D^tA_0)\Psi_n^t = 0. \tag{5.1.16}$$

由于这样的系数表示不唯一, 我们可能需要做一些多余的计算. 对于一个零项, 尽管它形式上非零, 并且参与计算, 但是它并不影响最后的计算结果, 因为在整个计算过程中只有加法和乘法, 所以在最后一步零项仍然是零. 因此上面的计算不依赖于等价系数的选取. 这里我们说x^t 的系数F_1^t 和F_2^t是等价的, 如果$\psi_n^t(F_1^t - F_2^t) = 0$, 即$(F_1^t - F_2^t)\Psi_n^t = 0$, 因为这时$F_1^tx^t = F_2^tx^t$.

注意到这个结果类似于一元多项式时的情形, 只是单变量的普通数乘被多变量的半张量积代替.

下面给出一个数值的例子.

例 5.1.16 设

$$\begin{aligned} P(x) &= 1 - (3x + 2y - z) + 2x^2 - 3xy + z^2 + 6x^2y - 2x^2z \\ &\quad + 7xy^2 - 9xyz + xz^2 - 2y^2z + y^3 + z^3, \end{aligned}$$

$$Q(x) := 1 - D(x) = 1 - (x + y - z),$$

其中$x = (x, y, z)^{\mathrm{T}} \in \mathbb{R}^3$, 那么

$$D = (1\ \ 1\ \ {-1}),$$

$$A_1 = (-3\ \ {-2}\ \ 1), \quad A_2 = (2\ \ {-3}\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 1),$$

$$A_3 = (0\ 6\ {-2}\ 0\ 7\ {-9}\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ {-2}\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1).$$

假设存在 $R(x)$, 使得 $Q(x)R(x) = P(x)$, 那么 $R(x) = 1 + B_1x + B_2x^2$. 根据(5.1.13), 我们计算 B_1 和 B_2 如下:

$$B_1 = A_1 + D = (-3\ \ {-2}\ \ 1) + (1\ \ 1\ \ {-1}) = (-2\ \ {-1}\ \ 0),$$

$$B_2 = A_2 + DB_1 = (0\ \ {-5}\ \ 2\ \ {-1}\ \ {-1}\ \ 1\ \ 0\ \ 0\ \ 1).$$

于是

$$\begin{aligned} DB_2 &= (0\ 0\ 0\ {-5}\ {-5}\ 5\ 2\ 2\ {-2}\ {-1}\ {-1}\ 1\ {-1}\ {-1}\ 1 \\ &\qquad 1\ 1\ {-1}\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ {-1}). \end{aligned}$$

显然$(A_3+DB_2)\Psi_3^3=0$, 于是我们得到

$$
\begin{aligned}
R(x) &= 1+B_1x+B_2x^2=1-(2x+y)\\
&\quad +(-6xy-y^2+2xz+yz+z^2).
\end{aligned}
$$

□

下面, 我们考虑高维多项式映射. 一个n 元k 次m 维多项式可以表示成

$$P(x)=A_0+A_1x+A_2x^2+\cdots+A_kx^k,\quad x\in\mathbb{R}^n,\ P(x)\in\mathbb{R}^m,$$

其中$A_j\in M_{m\times n^j}$, $j=0,\cdots,k$.

下面考虑两个多项式映射的张量积.

命题 5.1.17 设$P(x)=A_0+A_1x+A_2x^2+\cdots+A_px^p$, $Q(x)=B_0+B_1x+B_2x^2+\cdots+B_qx^q$, 那么

$$P(x)\ltimes Q(x)=\sum_{i=0}^{p+q}\sum_{j=0}^{i}A_j\otimes B_{i-j}x^i. \tag{5.1.17}$$

证明 回顾例2.3.12, 对于两个向量x 和y 有

$$(Ax)\ltimes(By)=(A\otimes B)(x\ltimes y).$$

利用这个公式以及半张量积的分配律就得到(5.1.17). □

下面的例子说明, 半张量积是处理具有多项式形式的数据的强大工具.

例 5.1.18 设$A(x)$ 是一个方阵, 它的元素都是t 次多项式, 并且$x\in\mathbb{R}^n$. $B(x)=I-Dx$ 是一个线性式, 这里$D\in M_{n\times n^2}$. 现在的问题是: 何时存在矩阵$C(x)$, 使得$A(x)=B(x)C(x)$? 我们想找到类似于命题5.1.15 中的多项式的一些性质.

首先, $A(x)$ 可以表示成

$$A(x)=A_0+A_1x+\cdots+A_tx^t. \tag{5.1.18}$$

这是因为, 根据次数可以很容易将每个元素中的k 次项提取出来, 这样$A(x)$ 可以表示成

$$A(x)=A_0+A^1+\cdots+A^t,$$

其中

$$A^k=\begin{bmatrix}a_{11}x^k&\cdots&a_{1n}x^k\\ \vdots&\ddots&\vdots\\ a_{n1}x^k&\cdots&a_{nn}x^k\end{bmatrix},\quad a_{ij}\in M_{1\times n^k},\quad k=1,2,\cdots,t.$$

上面的等式可以转化为

$$A^k = E_k(I_n \otimes x^k) = E_k \rtimes x^k, \tag{5.1.19}$$

其中

$$E_k = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in M_{n\times n^{k+1}}.$$

现在将(4.1.20) 应用到(5.1.19), 有

$$E_k \rtimes x^k = E_k \ltimes W_{[n^k,n]} \ltimes x^k \ltimes W_{[n,1]}.$$

根据定义, 可以直接验证

$$W_{[n,1]} = W_{[1,n]} = I_n. \tag{5.1.20}$$

利用(5.1.20), 并且令

$$A_k = E_k \ltimes W_{[n^k,n]}, \quad k = 1, \cdots, t.$$

就可以证明(5.1.18) 成立.

至于多项式情形, 我们假设存在$C(x) = C_0 + C_1 x + \cdots + C_{t-1}x^{t-1}$, 使得$A(x) = B(x)C(x)$. 比较

$$(I - Dx)(C_0 + C_1 x + \cdots + C_{t-1}x^{t-1}) = A_0 + A_1 x + \cdots + A_t x^t$$

两边的系数, 得到

$$\begin{cases} C_0 = A_0, \\ C_k x^k = A_k x^k + DxC_{k-1}x^{k-1}, \quad k = 1, \cdots, t. \end{cases} \tag{5.1.21}$$

要注意在上面的表示中所有的乘法都是半张量积. 利用(3.3.21), 有

$$xC_{k-1} = (I_n \otimes C_{k-1}) \ltimes x.$$

将它代入到(5.1.21), 并且利用多项式情形时同样的讨论, 不失一般性, 我们可以令

$$\begin{cases} C_0 = A_0, \\ C_k = A_k + D \rtimes C_{k-1}, \quad k = 1, \cdots, t-1. \end{cases} \tag{5.1.22}$$

递推迭代有

$$C_{t-1}=\sum_{i=1}^{t} D^{\rtimes(i-1)}\rtimes A_{t-i},$$

其中$D^{\rtimes i}=\underbrace{D\rtimes D\rtimes\cdots\rtimes D}_{i}$. 最后, 我们需要

$$Dx\ltimes C_{t-1}x^{t-1}=(D\rtimes C_{t-1})x^t+A_tx^t=0.$$

即

$$\left[\sum_{i=0}^{t} D^{\rtimes i}\rtimes A_{t-i}\right]x^t=0, \tag{5.1.23}$$

它可以保证$B(x)$ 是$A(x)$ 左除数.

一个类似的问题是: 设$B(x)=I-Dx$, 何时存在矩阵$C(x)$, 使得$A(x)=C(x)B(x)$? 同样, 类似于前面的讨论有

$$\left[\sum_{i=0}^{t} A_{t-i}\rtimes D^{\rtimes i}\right]x^t=0. \tag{5.1.24}$$

□

我们将上面的讨论总结如下:

定理 5.1.19 *设$A(x)$ 是一个方阵, 它的所有元素都是t 次多项式, 并且$x\in\mathbb{R}^n$, $B(x)=I-Dx$ 是一个线性式. 那么$B(x)$ 是$A(x)$ 的左因子, 当且仅当*

$$\left[\sum_{i=0}^{t} D^{\rtimes i}\rtimes A_{t-i}\right]\Psi_n^t=0. \tag{5.1.25}$$

$B(x)$ 是$A(x)$ 的右因子, 当且仅当

$$\left[\sum_{i=0}^{t} A_{t-i}\rtimes D^{\rtimes i}\right]\Psi_n^t=0. \tag{5.1.26}$$

证明 实际上(5.1.25) 和(5.1.26) 可以分别由(5.1.23) 和(5.1.24) 得到. 同多项式情形唯一的不同就是如何去乘对称化算子Ψ_n^t 的矩阵. 事实上, 对称化算子作用于矩阵的每一个元素. 也就是, 它要在$(x_{i,1},x_{i,n+1},x_{i,2n+1},\cdots)$ 上作用. 因此我们需要

$$\left[\sum_{i=0}^{t} D^{\rtimes i}\rtimes A_{t-i}\right](\Psi_n^t\otimes I_n)=0,$$

等等. 利用(3.1.3), 上式等价于(5.1.25). □

5.2 微 分 形 式

对于一个光滑函数$F(x):\mathbb{R}^n\to\mathbb{R}$, 我们定义它的微分为

$$DF(x)=\left(\frac{\partial F(x)}{\partial x_1},\cdots,\frac{\partial F(x)}{\partial x_n}\right),$$

定义它的梯度为

$$\nabla F(x)=(DF(x))^{\mathrm{T}}.$$

考虑矩阵$M(x)\in\mathcal{M}_{p\times q}$, 它的每个元素都是关于$x\in\mathbb{R}^n$ 的函数. 定义$M(x)$ 的微分 $DM(x)\in\mathcal{M}_{p\times nq}$ 为将$M(x)$ 的每个元素M_{ij} 替换为它的微分$\left(\frac{\partial M_{ij}}{\partial x_1},\cdots,\frac{\partial M_{ij}}{\partial x_n}\right)$, 即

$$DM(x)=\begin{bmatrix}\frac{\partial M_{11}(x)}{\partial x_1} & \cdots & \frac{\partial M_{11}(x)}{\partial x_n} & \cdots & \frac{\partial M_{1n}(x)}{\partial x_1} & \cdots & \frac{\partial M_{1n}(x)}{\partial x_n}\\ \vdots & & \vdots & & \vdots & & \vdots\\ \frac{\partial M_{n1}(x)}{\partial x_1} & \cdots & \frac{\partial M_{n1}(x)}{\partial x_n} & \cdots & \frac{\partial M_{nn}(x)}{\partial x_1} & \cdots & \frac{\partial M_{nn}(x)}{\partial x_n}\end{bmatrix}$$

同样, 定义$pn\times q$ 矩阵$M(x)$ 的梯度 $\nabla M(x)$ 为将$M(x)$ 的每个元素M_{ij} 替换为它的梯度$\left(\frac{\partial M_{ij}}{\partial x_1},\cdots,\frac{\partial M_{ij}}{\partial x_n}\right)^{\mathrm{T}}$.

容易证明, 矩阵的梯度可由它的微分经如下的块转置加转置实现.

$$\nabla M(x)=\left[(DM(x))^{\mathrm{T}(1\times n)}\right]^{\mathrm{T}}. \tag{5.2.1}$$

高阶的微分和梯度可以归纳地定义为

$$D^{k+1}M=D(D^kM)\in\mathcal{M}_{p\times n_{k+1}q},\quad k\geq 1,$$

$$\nabla^{k+1}M=\nabla(\nabla^kM)\in\mathcal{M}_{pn_{k+1}\times q},\quad k\geq 1.$$

下面的例子说明了引入这些记号的必要性.

例 5.2.1 在$\mathbb{R}^n$ 中, 我们称一个$n\times n$ 矩阵函数$M(x)$ 为一个结构矩阵. 利用它我们可以定义伪 Poisson 括号: $C^\infty(\mathbb{R}^n)\times C^\infty(\mathbb{R}^n)\to C^\infty(\mathbb{R}^n)$ 为

$$\{F(x),G(x)\}=dF(x)M(x)\nabla G(x),\quad F(x),G(x)\in C^\infty(\mathbb{R}^n). \tag{5.2.2}$$

定义由函数$F(x)$ 诱导的Hamilton 向量场为

$$X_F(x)=M(x)\nabla F(x),$$

$F(x)$ 称为 Hamilton 函数,

称下面的动态控制系统是一个伪Hamilton 控制系统:

$$\begin{cases} \dot{x} = X_F(x) + \sum\limits_{i=1}^{m} X_{G_i}(x)u_i(x), \quad x \in \mathbb{R}^n, \\ y = h(x) = (\{G_1, F\}, \{G_2, F\}, \cdots, \{G_m, F\})^{\mathrm{T}}. \end{cases} \tag{5.2.3}$$

如果我们想找一个光滑反馈$u_i(x)$, 使得闭环系统仍是一个保持原结构矩阵不变的伪Hamilton 系统, 即

$$\sum_{i=1}^{m} X_{G_i}(x)u_i(x) = M(x)\nabla\phi(x).$$

假设dG_i 线性无关, 那么我们可以找到$\{z_j(x),\ j = 1, \cdots, n-m\}$, 使得$\{dG_i,\ i = 1, \cdots, m,\ dz_j(x),\ j = 1, \cdots, n-m\}$ 线性无关. 根据隐函数定理, 我们可以局部表示$\phi(x) = \phi(G_i, z_j)$. 假设$M(x)$ 非奇异, 那么

$$\nabla\phi(x) = \sum_{i=1}^{m} \nabla G_i \frac{\partial\phi(G_i, z_j)}{\partial G_i} + \sum_{j=1}^{n-m} \nabla z_j \frac{\partial\phi(G_i, z_j)}{\partial z_j} = \sum_{i=1}^{m} \nabla G_i u_i.$$

因此

$$\begin{cases} \dfrac{\partial\phi(G_i, z_j)}{\partial z_j} = 0, \quad j = 1, \cdots, n-m, \\ u_i = \dfrac{\partial\phi(G_i, z_j)}{\partial G_i}, \quad i = 1, \cdots, m. \end{cases}$$

于是我们得出结论: 假设$M(x)$ 非奇异, 那么系统(5.2.2) 有一个伪Hamilton 实现, 当且仅当, 存在一个函数$\phi = \phi(G_i)$, 使得

$$u_i = \frac{\partial\phi(G_i)}{\partial G_i}, \quad i = 1, \cdots, m.$$

现在我们假设系统输出可检测[68], 即如果输出满足$\lim_{t\to\infty} h(x(t)) = 0$, 那么$\lim_{t\to\infty} x(t) = 0$. 于是我们可以假设Lyapunov 函数$L = h^{\mathrm{T}}(x)P(x)h(x)$, 其中$P(x) > 0$ (即正定). 对于闭环系统我们有

$$\dot{L} = D(h^{\mathrm{T}}(x)P(x)h(x))M(x)(\nabla F(x) + \nabla\phi(G(x))).$$

如果我们可以找到$P(x) > 0$ 和$\phi(G(x))$, 使得$\dot{L} < 0$, 则系统在原点是渐近稳定的.

□

从上面的例5.2.1 以及后面的许多讨论可以看出, 我们需要经常计算矩阵乘积的微分. 下面我们推导出一些矩阵函数乘积的导数公式. 首先, 我们考虑$D(A(x)\ltimes B(x))$, $A(x)\succ B(x)$.

命题 5.2.2 设$A(x)\succ_t B(x)$, $x\in\mathbb{R}^n$, 那么

$$D(A(x)\ltimes B(x)) = DA(x)\ltimes B(x) + A(x)\ltimes DB(x)\ltimes (I_s\otimes W_{[t,n]}), \tag{5.2.4}$$

其中s 表示$B(x)$ 的列数.

证明 不失一般性, 我们假设

$$A(x) = (a_{11},\cdots,a_{1t},\cdots,a_{q1},\cdots,a_{qt})$$

是一个行向量. 首先, 我们假设$s=1$, 即$B(x)=(b_1,\cdots,b_q)^{\mathrm{T}}$ 是一个列向量, 那么

$$A\ltimes B = \left(\sum_{k=1}^q a_{k1}b_k,\cdots,\sum_{k=1}^q a_{kt}b_k\right).$$

$$\begin{aligned}
&D(A\ltimes B) = \\
&\Bigg(\sum_{k=1}^q \frac{\partial a_{k1}}{\partial x_1}b_k + \sum_{k=1}^q a_{k1}\frac{\partial b_k}{\partial x_1},\cdots,\sum_{k=1}^q \frac{\partial a_{k1}}{\partial x_n}b_k + \sum_{k=1}^q a_{k1}\frac{\partial b_k}{\partial x_n},\cdots,\\
&\quad\sum_{k=1}^q \frac{\partial a_{kt}}{\partial x_1}b_k + \sum_{k=1}^q a_{kt}\frac{\partial b_k}{\partial x_1},\cdots,\sum_{k=1}^q \frac{\partial a_{kt}}{\partial x_n}b_k + \sum_{k=1}^q a_{kt}\frac{\partial b_k}{\partial x_n}\Bigg)\\
&= \left(\sum_{k=1}^q \frac{\partial a_{k1}}{\partial x_1}b_k,\cdots,\sum_{k=1}^q \frac{\partial a_{k1}}{\partial x_n}b_k,\cdots,\sum_{k=1}^q \frac{\partial a_{kt}}{\partial x_1}b_k,\cdots,\sum_{k=1}^q \frac{\partial a_{kt}}{\partial x_n}b_k\right)\\
&\quad+\left(\sum_{k=1}^q a_{k1}\frac{\partial b_k}{\partial x_1},\cdots,\sum_{k=1}^q a_{k1}\frac{\partial b_k}{\partial x_n},\cdots,\sum_{k=1}^q a_{kt}\frac{\partial b_k}{\partial x_1},\cdots,\sum_{k=1}^q a_{kt}\frac{\partial b_k}{\partial x_n}\right)\\
&:= I + II.
\end{aligned} \tag{5.2.5}$$

直接计算有

$$(DA)\ltimes B = \left(\sum_{k=1}^q \frac{\partial a_{k1}}{\partial x_1}b_k,\cdots,\sum_{k=1}^q \frac{\partial a_{k1}}{\partial x_n}b_k,\cdots,\sum_{k=1}^q \frac{\partial a_{kt}}{\partial x_1}b_k,\cdots,\sum_{k=1}^q \frac{\partial a_{kt}}{\partial x_n}b_k\right),$$

它就是(5.2.5) 的第一部分I.

同时, 直接计算可得

$$A\ltimes DB = \left(\sum_{k=1}^q a_{k1}\frac{\partial b_k}{\partial x_1},\cdots,\sum_{k=1}^q a_{kt}\frac{\partial b_k}{\partial x_1},\cdots,\sum_{k=1}^q a_{k1}\frac{\partial b_k}{\partial x_n},\cdots,\sum_{k=1}^q a_{kt}\frac{\partial b_k}{\partial x_n}\right).$$

现在$A \ltimes DB$ 和II 都由$\sum_{k=1}^{q} a_{ki}\frac{\partial b_k}{\partial x_j}$ 的元素组成. 但是在前者中它们是按照索引$Id(j,i;n,t)$ 排列的, 而在后者中它们是按照索引$Id(i,j;t,n)$ 排列的. 因此我们有

$$(II)^{\mathrm{T}} = W_{[n,t]}(A \ltimes DB)^{\mathrm{T}}.$$

于是

$$II = (A \ltimes DB)W_{[t,n]}.$$

至于一般情形, 当$s > 1$ 时, 记B 的第i 行为B_i, 那么

$$\begin{aligned}
&D(A \ltimes B)\\
&= (D(A \ltimes B_1), \cdots, D(A \ltimes B_s))\\
&= (DA \ltimes B_1 + (A \ltimes DB_1)W_{[n,t]}, \cdots, DA \ltimes B_s + (A \ltimes DB_s)W_{[n,t]})\\
&= (DA \ltimes B_1, \cdots, DA \ltimes B_s) + ((A \ltimes DB_1)W_{[n,t]}, \cdots, (A \ltimes DB_s)W_{[n,t]})\\
&= (DA \ltimes B) + ((A \ltimes DB_1), \cdots, (A \ltimes DB_s))(I_s \ltimes W_{[n,t]})\\
&= (DA \ltimes B) + (A \ltimes DB)(I_s \ltimes W_{[n,t]}).
\end{aligned}$$

于是(5.2.4) 成立. □

我们通过下面的例子说明上述公式的应用.

例 5.2.3 给定

$$A = \begin{bmatrix} x_1^2 & x_1x_2 & x_1x_2 & x_2^2 \\ x_2^2 & x_1x_2 & x_1x_2 & x_1^2 \end{bmatrix}, \quad B = \begin{bmatrix} \sin(x_1+x_2) & \cos(x_1+x_2) \\ -\cos(x_1+x_2) & \sin(x_1+x_2) \end{bmatrix}$$

那么

$$DA = \begin{bmatrix} 2x_1 & 0 & x_2 & x_1 & x_2 & x_1 & 0 & 2x_2 \\ 0 & 2x_2 & x_2 & x_1 & x_2 & x_1 & 2x_1 & 0 \end{bmatrix},$$

$$DB = \begin{bmatrix} \cos(x_1+x_2) & \cos(x_1+x_2) & -\sin(x_1+x_2) & -\sin(x_1+x_2) \\ \sin(x_1+x_2) & \sin(x_1+x_2) & \cos(x_1+x_2) & \cos(x_1+x_2) \end{bmatrix}.$$

为了记号的简便, 我们记$S=\sin(x_1+x_2)$, $C=\cos(x_1+x_2)$, 于是

$$
\begin{aligned}
&D(A(x)\ltimes B(x))\\
&=\begin{bmatrix}2x_1 & 0 & x_2 & x_1 & x_2 & x_1 & 0 & 2x_2\\ 0 & 2x_2 & x_2 & x_1 & x_2 & x_1 & 2x_1 & 0\end{bmatrix}\ltimes\begin{bmatrix}S & C\\ -C & S\end{bmatrix}\\
&\quad+\begin{bmatrix}x_1^2 & x_1x_2 & x_1x_2 & x_2^2\\ x_2^2 & x_1x_2 & x_1x_2 & x_1^2\end{bmatrix}\ltimes\begin{bmatrix}C & C & -S & -S\\ S & S & C & C\end{bmatrix}\ltimes\left(I_2\otimes W_{[2]}\right):=[a_{ij}],
\end{aligned}
$$

则

$$
\begin{aligned}
&a_{11}=2x_1S-x_2C+x_1^2C+x_1x_2S && a_{12}=-x_1C+x_1^2C+x_1x_2S\\
&a_{13}=x_2S+x_1x_2C+x_1^2S && a_{14}=x_1C+2x_2S+x_1x_2C+x_2^2S\\
&a_{15}=2x_1C+x_2S-x_1^2S+x_1x_2C && a_{16}=x_1S-x_1^2S+x_1x_2C\\
&a_{17}=x_2C-x_1x_2S+x_2^2C && a_{18}=x_1C+2x_2S-x_1x_2S+x_2^2C\\
&a_{21}=-x_2C+x_2^2C+x_1x_2S && a_{22}=2x_2S-x_1C+x_2^2C+x_1x_2S\\
&a_{23}=x_2S-2x_1C+x_1x_2C+x_1^2S && a_{24}=x_1S+x_1x_2C+x_1^2S\\
&a_{25}=x_1S-x_2^2S+x_1x_2C && a_{26}=2x_2C+x_1S-x_2^2S+x_1x_2C\\
&a_{27}=x_2C+2x_1S-x_1x_2S+x_1^2C && a_{28}=x_1C-x_1x_2S+x_1^2C.
\end{aligned}
$$

为了验证上面的结果, 直接计算有

$$
\begin{aligned}
&A(x)\ltimes B(x)\\
&=\begin{bmatrix}x_1^2S-x_1x_2C & x_1x_2S-x_2^2C & x_1^2C+x_1x_2S & x_1x_2C+x_2^2S\\ x_1^2S-x_1x_2C & x_1x_2S-x_1^2C & x_2^2C+x_1x_2S & x_1x_2C+x_2^2S\end{bmatrix}.
\end{aligned}
$$

微分它, 我们就有与上面一样的结果. □

将上述结果用于普通矩阵乘法, 则得到以下推论.

推论 5.2.4 对于普通矩阵乘法, 我们有

$$
D(A(x)B(x))=DA(x)\ltimes B(x)+A(x)DB(x). \tag{5.2.6}
$$

证明 注意到 $W_{[1,k]}=W_{[k,1]}=I_k$, $k>0$, 那么它就是 (5.2.4) 的一个特殊情形. □

下面考虑$A(x)\prec_t B(x)$, $t>1$ 时的情形. 我们可以利用公式$A(x)\ltimes B(x)=(A(x)\otimes I_t)B(x)$ 得到一个适当的形式.

首先, 我们给出两个引理, 它们本身也很有用.

引理 5.2.5 设A 是一个块行矩阵, 即

$$A=(A_{11},\cdots,A_{1n},\cdots,A_{k1},\cdots,A_{kn}),$$

这些块是由(i,j) 按照次序$Id(i,j;k,n)$ 排列的. 我们假设每块都有相同的维数, 那么

$$AW_{[n,k]}=(A_{11},\cdots,A_{k1},\cdots,A_{1n},\cdots,A_{kn}). \tag{5.2.7}$$

同理, 对于

$$B=\mathrm{col}(B_{11},\cdots,B_{1n},\cdots,B_{k1},\cdots,B_{kn}),$$

它的块也都有相同的维数, 那么

$$W_{[k,n]}B=\mathrm{col}(B_{11},\cdots,B_{k1},\cdots,B_{1n},\cdots,B_{kn}). \tag{5.2.8}$$

证明 设A_{ij} 的块列数是r, 容易看出$AW_{[n,k]}=A\ltimes W_{[n,k]}=A(W_{[n,k]}\otimes I_r)$. 显然$W_{[n,k]}$ 改变了A 中块的排列次序. 同理可证(5.2.8). □

引理 5.2.6 设$A(x)\in M_{p\times q}$, $x\in\mathbb{R}^n$, 那么

$$D(A\otimes I_k)=(DA\otimes I_k)(I_q\otimes W_{[k,n]}). \tag{5.2.9}$$

证明 容易验证$D(A\otimes I_k)$ 和$(DA\otimes I_k)$ 都可以被分割成$p\times q$ 个块, 分别将它们记为$D(A\otimes I_k)=E=\{E_{ij}\}$ 和$(DA\otimes I_k)=F=\{F_{ij}\}$, 那么

$$E_{ij}=D(a_{ij}I_k),\quad F_{ij}=da_{ij}\otimes I_k.$$

直接计算有, 如果F_{ij} 的列由(ij) 按照索引$Id(i,j;n,k)$ 排列, 那么E_{ij} 也是由同样的列组成, 但是它们按照索引$Id(j,i;k,n)$ 排列. 根据引理5.2.5 有, $E_{ij}=F_{ij}W_{[k,n]}$. 现在令

$$W=\mathrm{diag}(\underbrace{W_{[k,n]},\cdots,W_{[k,n]}}_{q}),$$

那么$E=FW$, 即(5.2.9) 成立. □

命题 5.2.7 设$A(x)\in M_{p\times q}$, $A(x)\prec_t B(x)$, $x\in\mathbb{R}^n$, 那么

$$D(A(x)\ltimes B(x))=DA(x)\ltimes(I_q\otimes W_{[t,n]})\ltimes B(x)+A\ltimes DB(x). \tag{5.2.10}$$

证明 由于$A(x)\ltimes B(x)=(A(x)\otimes I_t)B(x)$, 并利用(5.2.6) 和(5.2.9) 有

$$\begin{aligned}D(A(x)\ltimes B(x)) &= D[(A(x)\otimes I_t)B(x)]\\ &= (DA\otimes I_t)(I_q\otimes W_{[t,n]})\ltimes B(x)+A(x)\ltimes DB(x),\end{aligned}$$

于是(5.2.10) 成立. □

下面给出一个用半张量积和微分算子对向量场李导数进行计算的例子.

例 5.2.8 设$f(x), g(x) \in V(M)$ 是M 上的两个向量场. 那么$g(x)$ 关于$f(x)$ 的李导数在局部可以表示为[68]

$$\mathrm{ad}_f(g) = [f, g] = Dgf - Dfg. \tag{5.2.11}$$

考虑二阶导数$\mathrm{ad}_f^2(g)$. 利用(5.2.4) 有

$$\begin{aligned}\mathrm{ad}_f^2(g) &= D(Dgf - Dfg)f - Df(Dgf - Dfg)\\ &= D^2 g \ltimes f^2 + DgDff - D^2 f \ltimes g \ltimes f - 2DfDgf + (Df)^2 g.\end{aligned}$$

□

下面, 我们考虑乘积的高阶导数. 为此, 我们需要做一些准备工作. 定义$i+1$ 重指标集为

$$S(i,k) = \left\{ d = (d_1, \cdots, d_{i+1}) \in \mathbb{Z}_+^{i+1} \;\middle|\; \sum_{j=1}^{i+1} d_j = k \right\}.$$

如果$d = (d_1, \cdots, d_{i+1}) \in S(i,k)$, 我们定义微分$D_d(B)$ 为

$$D_d B = D[(\cdots D(D(B \otimes I_{n^{d_1}}) \otimes I_{n^{d_2}}) \cdots \otimes I_{n^{d_i}}) \otimes I_{n^{d_{i+1}}}]. \tag{5.2.12}$$

并且记

$$D^{S(i,k)} B = \sum_{d \in S(i,k)} D_d B. \tag{5.2.13}$$

有了这些记号, 我们有

推论 5.2.9 对于普通矩阵乘法, 有

$$D^k(A(x)B(x)) = \sum_{j=0}^{k} D^j A(x) D^{S(k-j,j)}(B(x)). \tag{5.2.14}$$

证明 我们用数学归纳法证明. 当$k=1$ 时, 它就是(5.2.6). 假设(5.2.14) 对k 成立. 考虑$D^{k+1}(A(x)B(x))$ 中$D^{j+1}AD^{S(k-j,j+1)}B$ 的项, 它们来自于$D^k(A(x)B(x))$ 中的两组项: $D^j AD^{S(k-j,j)}B$ 和$D^{j+1}AD^{S(k-j-1,j+1)}B$. 我们微分第一组中的$A(x)$ 和第二组中的$B(x)$, 就得到所需要的项. 于是问题变为我们只要证明

$$D^{S(k-j,j)} B \otimes I_n + D[D^{S(k-j-1,j+1)} B] = D^{S(k-j,j+1)} B. \tag{5.2.15}$$

注意到, 对于

$$d=(d_1,\cdots,d_{k-j},d_{k-j+1}+1),\quad d_1+\cdots+d_{k-j+1}=j,$$

左边的第一组包含所有的D_d, 同时, 对于

$$d=(d_1,\cdots,d_{k-j},0),\quad d_1+\cdots+d_{k-j}=j+1,$$

左边的第二组也包含所有的D_d. 这两个d 的不相交组构成了$S(k-j,j+1)$ 的集合. 因此, (5.2.14) 成立. □

例 5.2.10　设$k=3$. 我们计算导数

$$\begin{cases} D^{S(3,0)}B=D^3B,\\ D^{S(2,1)}B=D^2(B\otimes I_n)+D(DB\otimes I_n)+D^2(B)\otimes I_n,\\ D^{S(1,2)}B=D(B\otimes I_{n^2})+D(B\otimes I_n)\otimes I_n+DB\otimes I_{n^2},\\ D^{S(0,3)}B=D^3B\otimes I_{n^3}. \end{cases}$$

因此

$$\begin{aligned} D^3(AB) \;=\;& AD^3B\\ &+DA\ltimes(D^2(B\otimes I_n)+D(DB\otimes I_n)+D^2(B)\otimes I_n)\\ &+D^2A\ltimes(D(B\otimes I_{n^2})+D(B\otimes I_n)\otimes I_n+DB\otimes I_{n^2})\\ &+D^3A(B\otimes I_{n^3}). \end{aligned}$$

□

下面, 我们考虑梯度.

引理 5.2.11　(1) 设$x\in\mathbb{R}^n$ 是一个列向量, 那么

$$x^{\mathrm{T}}=V_r^{\mathrm{T}}(I_n)\rtimes x. \tag{5.2.16}$$

(2) 设$x\in\mathbb{R}^n$ 是一个行向量, 那么

$$x^{\mathrm{T}}=x\ltimes V_r(I_n). \tag{5.2.17}$$

我们将引理的验证留给读者. 利用这个命题可以得到下述结论.

命题 5.2.12 设$A(x) \in M_{p\times q}$, 其中$x \in \mathbb{R}^n$, 那么

$$DA(x) = (I_p \otimes V_r^{\mathrm{T}}(I_n)) \rtimes \nabla A(x), \tag{5.2.18}$$

反之则有

$$\nabla A(x) = DA(x) \ltimes (I_q \otimes V_r(I_n)). \tag{5.2.19}$$

证明 将$DA(x)$ 和$\nabla A(x)$ 按照自然方式分割成$p \times q$ 块. 实际上, 我们只需将每块从列形式转化为行形式, 或者反之. 分别利用(5.2.16) 和(5.2.17), 以及分块乘法技巧就可以证明命题成立. □

本节的最后, 我们给出乘积梯度的公式.

命题 5.2.13 给定两个矩阵函数$A(x), B(x)$, 其中$x \in \mathbb{R}^n$.

(1) 如果$A(x) \prec B(x)$, 那么

$$\nabla(A(x)B(x)) = (\nabla A(x))B(x) + A(x)(\nabla B(x)). \tag{5.2.20}$$

(2) 如果$A(x) \succ_t B(x)$, 设$A(x) \in M_{m\times tp}$, $B(x) \in M_{p\times q}$, 那么

$$\nabla(A(x)B(x)) = (\nabla A(x))B(x) + A(x)DB(x)[I_{qt} \otimes V_r(I_n)]. \tag{5.2.21}$$

证明 利用左半张量积分块乘法的性质, 可以直接验证(5.2.20) 成立. 至于(5.2.21), 第一项可以直接通过分块乘法得到. 对于第二项, 假设$A(x) = A$为常值, 既然$A \succ B(x)$, 那么

$$\nabla(AB(x)) = D(AB(x)) \ltimes (I_{qt} \otimes V_r(I_n)) = ADB(x)[I_{qt} \otimes V_r(I_n)].$$

□

5.3 基变换

我们已经知道x^k 是B_n^k 的一组冗余基底, 并且(5.1.1) 中的F 不唯一. B_n^k 的自然基底, 记为N_n^k, 是

$$N_n^k := \left\{ x_1^{d_1} \cdots x_n^{d_n} \,\middle|\, \sum_{j=1}^n d_j = k \right\}.$$

为方便起见, 也用N_n^k 来表示按照字母序排列的有序基底元素矩阵. 即

$$x_1^{d_1} \ \cdots \ x_n^{d_n} \prec x_1^{t_1} \ \cdots \ x_n^{t_n}.$$

如果存在一个指标$i \le n$ 使得$d_k = t_k$, $k < i$ 且$d_i > t_i$.

简记$Id(I;n^k) := Id(i_1,\cdots,i_k;\underbrace{n,\cdots,n}_{k})$. 容易看出元素

$$x_{i_1} \cdots x_{i_k} \in x^k$$

是按照次序$Id(I;n^k)$ 排列的.

定义$Id(I;n^k)$ 的子指标集

$$Is(I;n^k) = \{(i_1,\cdots,i_k) \in Id(I;n^k) \,|\, i_1 \le i_2 \le \cdots \le i_k\}$$

为对称指标集. 注意到对于一个对称形式, 对称指标只是忽略重复元素, 索引不同元素. 例如, $x_1x_1x_2$, $x_1x_2x_1$, $x_2x_1x_1$ 都属于x_2^3, 或者说(112) (121) (211) $\in Id(I;2^3)$, 但是只有(112) $\in Is(I;2^3)$.

回顾指标$I = (i_1,\cdots,i_k) \in Id(I;n^k)$ 的数量$C_I \in \mathbb{Z}_+^n$ 的定义: $C_I(k)$ 就是整数k 在I 中出现的次数. 例如, 考虑$I = (11424) \in Id(I;6^5)$, 我们有$C_I = (210200)$.

这样我们定义映射$\xi : Id(I;n^k) \to \mathbb{Z}_+^n$, $I \mapsto C_I$.

由定义直接得到下面的命题:

命题 5.3.1　(1) 定义ξ 的域为

$$R(n^k) = \left\{ (c_1,\cdots,c_n) \in \mathbb{Z}_+^n \;\middle|\; \sum_{j=1}^n c_j = k \right\}. \tag{5.3.1}$$

(2) 如果ξ 限制在$Is(I;n^k)$ 上, 那么$\xi|_{Is(I;n^k)} : Is(I;n^k) \to R(n^k)$ 是一个双射.

(3) 定义$R(n^k)$ 上的次序$\prec$ 为

$$(c_1,\cdots,c_n) \prec (d_1,\cdots,d_n).$$

如果存在一个$s \le n$ 使得$c_1 = d_1$, $\cdots$, $c_{s-1} = d_{s-1}$ 和$c_s > d_s$.

在这个次序下$\xi : Is(I;n^k) \to R(n^k)$ 是一个保序映射.

注意到$R(n^k)$ 是指标集, 它索引自然基底N_n^k. 例如, $I = (11335) \in Id(I;6^5)$, 那么它索引$B_n^k = B_6^5$ 中的元素为$x_1x_1x_3x_3x_5$. 即使它属于$Is(I;6^5)$, 使用这样的元素作为基底也不方便. 相应地, $C_I = (2,0,2,0,1,0)$. 我们用它索引N_n^k 中的元素$x_1^2x_2^0x_3^2x_4^0x_5^2x_6^0 = x_1^2x_3^2x^5$. 因此我们记$N_n^K$ 的基底为

$$\left\{ x^{C_I} = \prod_{j=1}^n x_j^{c_j} \;\middle|\; I \in Is(I;n^k) \right\}.$$

命题 5.3.2 $Is(I;n^k)$ 的基数(大小) 是

$$|Is(I;n^k)| = \frac{(n+k-1)!}{k!(n-1)!}, \quad k \geq 0,\ n \geq 1. \tag{5.3.2}$$

证明 由于$Ct|_{Is(I;n^k)} : Is(I;n^k) \to R(n^k)$ 是一个双射, 所以我们只需要考虑

$$R(n^k) = \left\{ (c_1, \cdots, c_n) \in \mathbb{Z}_+^n \,\middle|\, \sum_{j=1}^{n} c_j = k \right\}$$

的大小. 让c_1 从0 跑到k, 那么显然有

$$|R(n^k)| = \sum_{j=0}^{k} \left| R((n-1)^j) \right|.$$

利用等式

$$\binom{n-1}{0} + \binom{n}{1} + \cdots + \binom{n+k-1}{k} = \binom{n+k}{k}, \tag{5.3.3}$$

由数学归纳法就可以证明结论. □

有趣的是, $Is(I;n^k)$ 的大小可以很容易从下表得到, 这个表除第一行和第一列的元素全都是1 外, 其他元素都是由它上边和左边的元素的和，见表5.3.1.

表 5.3.1 $Is(I;n^k)$ 的大小

$k\backslash \dim n$	1	2	3	4	5	6	$\cdots$
0	1	1	1	1	1	1	
1	1	2	3	4	5		
2	1	3	6	10			
3	1	4	10				
4	1	5					
5	1						
$\vdots$							

下面构造两个矩阵: $T_B(n,k)$ 和$T_N(n,k)$.

设$s = |Is(I;n^k)|$, $t = n^k$, 那么$T_B(n,k) \in M_{s\times t}$, $T_N(n,k) \in M_{t\times s}$. 这两个矩阵是用来将一种基底变换为另一种基底的.

首先, 我们构造矩阵$T_B(n,k)$ 如下:

第 1 步: 将一个$s\times t$ 矩阵的列按照次序$Id(J;n^k)$ 排列, 并且它的行按照次序$Is(I;n^k)$ 排列.

第 2 步: 按如下方式给每一行赋值:

假设行指标是$I=(i_1,i_2,\cdots,i_k)\in Is(I;n^k)$, 其中$i_1\le i_2\le\cdots\le i_k$. 对于这行中的任意元素, 如果这个元素的列指标$J\in P_I$, 即$J=(i_{\sigma(1)},i_{\sigma(2)},\cdots,i_{\sigma(k)})$, 对于某个$\sigma\in S_k$, 即这个元素的列指标是它的行指标的置换, 那么给这个元素赋值为

$$\alpha_I=\frac{C_I!}{k!}=\frac{\prod_{j=1}^n C_I(j)!}{k!},$$

否则令它为0. 这样就得到矩阵$T_B(n,k)$. 即, $T_B(n,k)$ 的元素$\beta_{I,J}$ 为

$$\beta_{I,J}=\begin{cases}\alpha_I, & J\in P_I,\\ 0, & \text{其他}.\end{cases}$$

接着, 构造矩阵$T_N(n,k)$:

第 1 步. 设T 是一个$t\times s$ 矩阵. 它的行按照次序$Id(J;n^k)$ 排列, 它的列按照次序$Is(I;n^k)$ 排列.

第 2 步. 按如下方式给它的每列赋值:

假设这列的指标是$J=(j_1,j_2,\cdots,j_k)\in Is(J;n^k)$. 于是对于这行的一个元素, 如果它的行指标$I\in P_J$, 那么将这个元素赋值为1, 否则赋值为0.

这样就得到矩阵$T_N(n,k)$, 即其元素$n_{I,J}$ 为

$$n_{I,J}=\begin{cases}1, & I\in P_J,\\ 0, & \text{其他}.\end{cases}$$

简记 N_n^k 的基底为 $x_{(k)}$. 例如, 当 $n=3$ 时, 可以将 $x_{(2)}$ 表示成 $x_{(2)}=(x_1^2,\ x_1x_2,\ x_1x_3,\ x_2^2,\ x_2x_3,\ x_3^2)^{\mathrm{T}}$.

下面的命题是上面定义的直接推论.

命题 5.3.3 (1) $x_{(k)}$ 和x^k 有如下关系

$$\begin{cases}x_{(k)}=T_B(n,k)x^k,\\ x^k=T_N(n,k)x_{(k)}.\end{cases}\tag{5.3.4}$$

(2) 设$p(x)\in B_n^k$ 是一个k 次齐次多项式, 且$p(x)=Fx^k=Sx_{(k)}$, 那么

$$S=FT_N(n,k).\tag{5.3.5}$$

同时,$ST_B(n,k)$ 是F 的一个对称表示. 特别, 如果F 对称, 那么$F=ST_B(n,k)$.

例 5.3.4 设$n=2$, $k=3$.

(1) 构造$T_B(n,k)$ 为

$$T_B(2,3) = \begin{array}{c} \begin{array}{cccccccc} (111) & (112) & (121) & (122) & (211) & (212) & (221) & (222) \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 0 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/3 & 0 & 1/3 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} (111) \\ (112) \\ (122) \\ (222) \end{array}. \tag{5.3.6}$$

同理, 我们有

$$T_N(2,3) = \begin{array}{c} \begin{array}{cccc} (111) & (112) & (122) & (222) \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \begin{array}{c} (111) \\ (112) \\ (121) \\ (122) \\ (211) \\ (212) \\ (221) \\ (222) \end{array} \tag{5.3.7}$$

(2) 设$f(x)=x_1^3+3x_1^2x_2-x_1x_2^2+x_2^3$, 将它重写为$f(x)=(1\ 3\ -1\ 1)x_3$. 利用(5.3.6), 有

$$f(x)=(1\ 3\ -1\ 1)T_B(n,k)x^k=\left(1\ 1\ 1\ -\frac{1}{3}\ 1\ -\frac{1}{3}\ -\frac{1}{3}\ 1\right)x^3.$$

(3) 设$f(x)=(1\ 2\ 1\ 1\ -1\ -1\ -2\ -1)x^3$, 利用(5.3.7), 有

$$f(x)=(1\ 2\ 1\ 1\ -1\ -1\ -2\ -1)T_N(n,k)x_{(k)}=(1\ 2\ -2\ -1)x_{(3)}.$$

于是得到$f(x)$ 的对称表示为

$$\begin{aligned}(1\ 2\ -2\ -1)x_3 &= (1\ 2\ -2\ -1)T_B(2,3)x^3 \\ &= \left(1\ \frac{2}{3}\ \frac{2}{3}\ -\frac{2}{3}\ \frac{2}{3}\ -\frac{2}{3}\ -\frac{2}{3}\ -1\right)x^3.\end{aligned}$$

□

5.4 多维映射的Taylor 展开

矩阵张量积的微分形式为我们提供了表示Taylor 级数的简便方法.

定理 5.4.1 (Taylor 级数) 设$F : \mathbb{R}^m \to \mathbb{R}^n$ 是一个解析映射, 那么它的Taylor 展开为

$$F(x) = F(x_0) + \sum_{k=1}^{\infty} \frac{1}{k!} D^k F(x_0) \ltimes (x - x_0)^k. \tag{5.4.1}$$

在文献[1] 中也有类似的表示形式. 这个公式的好处是和单变量函数的表示形式一致, 并且每一项中的乘积满足结合律.

接下来, 我们考虑一个局部微分同胚的逆映射. 我们要找到这个逆映射的Taylor 展开.

设 $F : \mathbb{R}^n \to \mathbb{R}^n$ 是一个解析映射. 不失一般性, 可令 $F(0) = 0$, 其余的 F 用 $F - F(0)$ 替换. 利用 Taylor 展开, 我们可以将它简单地表示成

$$y = F_1 x + F_2 x^2 + F_3 x^3 + \cdots, \tag{5.4.2}$$

其中

$$F_k = \frac{1}{k!} D^k F|_0, \quad k \geq 1.$$

利用(5.1.17), 我们有

$$\begin{bmatrix} y \\ y^2 \\ \vdots \\ y^k \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & \cdots & A_{1k} \\ 0 & A_{22} & A_{23} & \cdots & A_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & A_{kk} \end{bmatrix} \begin{bmatrix} x \\ x^2 \\ \vdots \\ x^k \end{bmatrix} + O(\|x\|^{k+1}), \tag{5.4.3}$$

其中$A_{1i} = F_i$, $i = 1, 2, \cdots$; A_{ki}, $k > 1$ 可由下式计算得到:

$$A_{ki} = \sum_{j_1 + j_2 + \cdots + j_k = i} (F_{j_1} \otimes F_{j_2} \otimes \cdots \otimes F_{j_k}), \quad i \geq k. \tag{5.4.4}$$

设$y = F(x)$ 是一个微分同胚, 那么$F_1 = A_{11} = J_F$ 是F 的Jacobi 矩阵, 并且它是可逆的. 而且由于$A_{kk} = \underbrace{F_1 \otimes \cdots \otimes F_1}_{k}$, 所以它也可逆. 因此我们由(5.4.3) 可以得到

$$\begin{bmatrix} x \\ x^2 \\ \vdots \\ x^k \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} & B_{13} & \cdots & B_{1k} \\ 0 & B_{22} & B_{23} & \cdots & B_{2k} \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & B_{kk} \end{bmatrix} \begin{bmatrix} y \\ y^2 \\ \vdots \\ y^k \end{bmatrix} + R_{k+1}. \tag{5.4.5}$$

记(5.4.5) 右边的系数矩阵为B^{kk}, 那么可以归纳地定义B^{kk} 为

$$\begin{cases} B^{11} = B_{11} = F_1^{-1}, \\ B^{t+1,t+1} = \begin{bmatrix} B^{tt} & -B^{tt}A^{t,t+1}A_{t+1,t+1}^{-1} \\ 0 & A_{t+1,t+1}^{-1} \end{bmatrix}, \quad t \geq 1, \end{cases} \tag{5.4.6}$$

其中

$$A_{t+1,t+1}^{-1} = \underbrace{F_1^{-1} \otimes \cdots \otimes F_1^{-1}}_{t+1}, \quad A^{t,t+1} = \begin{bmatrix} A_{1,t+1} \\ \vdots \\ A_{t,t+1} \end{bmatrix}.$$

定理 5.4.2 设$y = F(x) : \mathbb{R}^n \to \mathbb{R}^n$, $F(0) = 0$ 是一个局部微分同胚, 那么它的逆映射$x = F^{-1}y$ 的Taylor 展开为

$$x = B_{11}y + B_{12}y^2 + \cdots + B_{1k}y^k + O(\|y\|^{k+1}),$$

其中系数B_{1k} 是由(5.4.6) 所示.

证明 这个表达式实际上就是前面讨论的总结. 我们需要提到的就是(5.4.5) 中的余项R_{k+1}. 注意到对于微分同胚$y = F(x)$, $y(0) = 0$, 我们有$O(\|x\|^k) = O(\|y\|^k)$, 也就意味着$R_{k+1} = O(\|y\|^{k+1})$. □

在实际使用时可以很方便地将多变量映射的Taylor 展开表示在自然基底上, 这样没有多余的项. 为此, 我们定义两个矩阵:

$$\begin{cases} T^N(n,k) = \mathrm{diag}(I_n, T_N(n,2), T_N(n,3), \cdots, T_N(n,k)), \\ T^B(n,k) = \mathrm{diag}(I_n, T_B(n,2), T_B(n,3), \cdots, T_B(n,k)). \end{cases}$$

由$T_B(n,k)$ 和$T_N(n,k)$ 的性质容易看出, 如果

$$\begin{bmatrix} x \\ x^2 \\ \vdots \\ x^k \end{bmatrix} = B^{kk} \begin{bmatrix} Y \\ Y^2 \\ \vdots \\ Y^k \end{bmatrix},$$

那么,在自然基底下它变成

$$\begin{bmatrix} x \\ x_{(2)} \\ \vdots \\ x_{(k)} \end{bmatrix} = \mathrm{T}^B(n,k)B^{kk}\mathrm{T}^N(n,k)\begin{bmatrix} Y \\ Y_{(2)} \\ \vdots \\ Y_{(k)} \end{bmatrix} \tag{5.4.7}$$

例 5.4.3　考虑映射$y = F(x)$

$$\begin{cases} y_1 = \sin(x_1) + x_2 - x_2^2, \\ y_2 = \log(1 + x_1 - x_2). \end{cases} \tag{5.4.8}$$

通过Taylor 展开, (5.4.8) 可以表示成

$$\begin{cases} y_1 = x_1 + x_2 - x_2^2 - \frac{1}{6}x_1^3 + O(\|x\|^4), \\ \\ y_2 = x_1 - x_2 - \frac{1}{2}(x_1 - x_2)^2 + \frac{1}{3}(x_1 - x_2)^3 + O(\|x\|^4). \end{cases} \tag{5.4.9}$$

直接计算有表5.4.1的形式.

表 5.4.1　展开系数

y\\x	x_1	x_2	x_1^2	x_1x_2	x_2^2	x_1^3	$x_1^2x_2$	$x_1x_2^2$	x_2^3	$\cdots$
y_1	1	1	-1	0	0	-1/6	0	0	0	$\cdots$
y_2	1	-1	-1/2	1	-1/2	1/3	-1	1	-1/3	$\cdots$
y_1^2	0	0	1	2	1	0	-2	0	-2	$\cdots$
y_1y_2	0	0	1	0	-1	-1/2	1/2	1/2	-1/2	$\cdots$
y_2^2	0	0	1	-2	1	-1	3	-3	1	$\cdots$
y_1^3	0	0	0	0	0	1	3	3	1	$\cdots$
$y_1^2y_2$	0	0	0	0	0	1	1	-1	-1	$\cdots$
$y_1y_2^2$	0	0	0	0	0	1	-1	-1	1	$\cdots$
y_2^3	0	0	0	0	0	1	-3	3	1	$\cdots$
$\vdots$										

那么我们可以得到它的逆矩阵为:

表 5.4.2 逆映射展开系数

x\y	y_1	y_2	y_1^2	y_1y_2	y_2^2	y_1^3	$y_1^2y_2$	$y_1y_2^2$	y_2^3
x_1	0.5	0.5	0.125	0.25	0.375	0.0208	0.125	0.0625	0.2083
x_2	0.5	-0.5	0.125	0.25	-0.125	0.0764	-0.0417	0.2294	-0.0139
x_1^2	0	0	0.25	0.5	0.25	0.0833	0	0.25	0.1667
x_1x_2	0	0	0.25	0	-0.25	0.25	-0.5	0.5	-0.5
x_2^2	0	0	0.25	-0.5	0.25	0.0833	0	-0.25	0.1667
x_1^3	0	0	0	0	0	0.0833	0.5	0.25	0.1667
$x_1^2x_2$	0	0	0	0	0	0.1667	0	0	-0.1667
$x_1x_2^2$	0	0	0	0	0	0.0833	0	-0.25	0.1667
x_2^3	0	0	0	0	0	0.1667	-0.5	0.5	-0.1667
$\vdots$									

从表5.4.2 矩阵的第一块行, 我们可以得到(5.4.8) 中定义的映射$y = F(x)$ 的逆映射是

$$
\begin{aligned}
x &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = F^{-1}(y) \\
&= \begin{bmatrix} 0.5y_1 + 0.5y_2 + 0.125y_1^2 + 0.25y_1y_2 + 0.375y_2^2 \\ +0.0208y_1^3 + 0.125y_1^2y_2 + 0.0625y_1y_2^2 + 0.2083y_2^3 \\ \\ 0.5y_1 - 0.5y_2 + 0.125y_1^2 + 0.25y_1y_2 - 0.125y_2^2 \\ +0.764y_1^3 - 0.0417y_1^2y_2 + 0.2292y_1y_2^2 - 0.0139y_2^3 \end{bmatrix} + O(\|y\|^4).
\end{aligned}
$$

□

5.5 基本微分公式

本章最后的两节将讨论向量场的计算问题, 它们在非线性系统的几何理论中起着关键作用. 为避免讨论可微分次数及Taylor 展开式的收敛性, 我们假定M, N 等为解析流形. $C^\omega(M)$, $V^\omega(M)$ 及$V^{*\omega}(M)$ 分别为流形上的解析函数集合, 解析向量场集合及解析余向量场集合. 我们只讨论这些对象.

设$h(x) \in C^\omega(M)$. 利用Taylor 展式可得

$$h(x) = h_0 + h_1 x + h_2 x^2 + \cdots ,$$

这里h_k 为$1\times n^k$ 定常行向量. 于是, $h(x)$ 的微分可表示为

$$Dh(x)=h_1+h_2D(x^2)+h_3D(x^3)+\cdots. \tag{5.5.1}$$

类似地, 令$X(x)\in V^{\omega}(M)$. 利用Taylor 展式, 我们有

$$X(x)=X_0+X_1x+X_2x^2+\cdots,$$

这里X_k 为$n\times n^k$ 定常矩阵. 于是$X(x)$ 的Jacobi 矩阵可表示为

$$J_X(x)=X_1+X_2D(x^2)+X_3D(x^3)+\cdots. \tag{5.5.2}$$

与此类似, 我们很容易发现, 在几何计算中计算高阶多项式的微分$D(x^k)$ 是关键. 本节给出它的计算公式.

引理 5.5.1

$$\begin{aligned} D(x^k)= &\ W_{[n^{k-1},n]}x^{k-1}+xW_{[n^{k-2},n]}x^{k-2}+\cdots\\ &+x^{k-2}W_{[n,n]}x+x^{k-1}\otimes I_n, \qquad k\ge 2. \end{aligned} \tag{5.5.3}$$

证明 利用数学归纳法. 显然

$$Dx=I_n.$$

利用(5.2.4), 我们有

$$\begin{aligned} D(x^2) &= Dx\ltimes(1\otimes W_{[n,n]})\ltimes x+x\ltimes I_n\\ &= I_n\ltimes W_{[n,n]}\ltimes x+(x\otimes I_n)I_n=W_{[n,n]}\ltimes x+x\otimes I_n. \end{aligned}$$

设(5.5.3) 对k 成立. 根据(5.2.9)

$$D(x\otimes I_{n^k})=(I_n\otimes \dot{I}_{n^k})(1\otimes W_{[n^k,n]})=I_{n^{k+1}}W_{[n^k,n]}=W_{[n^k,n]}.$$

于是

$$\begin{aligned} D(x^{k+1}) &= D[(x\otimes I_{n^k})x^k]=D(x\otimes I_{n^k})x^k+(x\otimes I_{n^k})D(x^k)\\ &= W_{[n^k,n]}x^k+(x\otimes I_{n^k})(W_{[n^{k-1},n]}x^{k-1}\\ &\quad +xW_{[n^{k-2},n]}x^{k-2}+\cdots+x^{k-1}\otimes I_n)\\ &= W_{[n^k,n]}x^k+xW_{[n^{k-1},n]}x^{k-1}+\cdots+x^k\otimes I_n. \end{aligned}$$

□

下面的定理给出微分计算的基本公式.

定理 5.5.2 设$x \in \mathbb{R}^n$, 那么

$$D(x^{k+1}) = \Phi_k^n x^k, \quad k \geq 0, \tag{5.5.4}$$

这里

$$\Phi_k^n = \sum_{s=0}^{k} I_{n^s} \otimes W_{[n^{k-s},n]}. \tag{5.5.5}$$

证明 利用引理5.5.1 以及下面的换序公式

$$x^p W_{[n^s,n]} = (I_{n^p} \otimes W_{[n^s,n]})x^p,$$

立即可推出(5.5.4). □

注 由于$I_1 = 1$ 是一个数, 而$W[1,n] = I_n$, 于是可知

$$\Phi_0^n = I_n.$$

当不会产生混淆时, 我们把Φ_k^n 简记为Φ_k.

5.6 李导数

本节讨论在非线性控制理论中用到的几种李导数及其在局部坐标下的计算公式.

定义 5.6.1 设$F: M \to N$ 为一微分同胚.

(1) 对一个函数$h(x) \in C^\omega(N)$, F 的导出映射$F^*: C^\omega(N) \to C^\omega(M)$ 定义为

$$F^*(h) = h \circ F \in C^\omega(M).$$

(2) 对一个向量场$X \in V^\omega(M)$, F 的导出映射$F_*: V^\omega(M) \to V^\omega(N)$ 定义为

$$F_*(X)(h) = X(h \circ F), \quad \forall h \in C^\omega(N).$$

(3) 对一个余向量场$\alpha \in V^{*\omega}(N)$, F 的导出映射$F^*: V^{*\omega}(N) \to V^{*\omega}(M)$ 定义为

$$\langle F^*(\alpha), X \rangle = \langle \alpha, F_*(X) \rangle, \quad \forall X \in V^r(M).$$

如果F 是一个局部微分同胚, 以上的各种映射均为局部定义的.

对于一个向量场$X \in V^{\omega}(M)$, 它的以$x(0) = x_0$ 为初值的积分曲线记作$\phi_t^X(x_0)$. 为方便计, 我们假定X 是完备的, 即它的积分曲线对所有的$t \in \mathbb{R}$ 均有定义. 那么对每一个固定的t, $\phi_t^X : M \to M$ 是一个微分同胚[10].

定义 5.6.2　设$X \in V^{\omega}(M)$, $h \in C^{\omega}(M)$. 那么h 对X 的李导数, 记作$L_X(h)$, 定义为

$$L_X(h) = \lim_{t\to 0} \frac{1}{t}\left[(\phi_t^X)^* f(x) - f(x)\right]. \tag{5.6.1}$$

命题 5.6.3　在局部坐标下李导数(5.6.1) 可表示为

$$L_X(h) = \langle dh, X\rangle = \sum_{i=1}^{n} X_i \frac{\partial h}{\partial x_i}. \tag{5.6.2}$$

证明　根据定义$(\phi_t^X)^* h(x) = h(\phi_t^X(x))$. 因此它对$t$ 的Taylor 展式为

$$h\left(\phi_t^X(x)\right) = h(x) + tdf \cdot X(x) + O(t^2)$$

将它代入(5.6.1) 即得(5.6.2). □

定义 5.6.4　设X, $Y \in V(M)$. 则Y 对X 的李导数, 记作$\mathrm{ad}_X(Y)$, 定义为

$$\mathrm{ad}_X(Y) = \lim_{t\to 0} \frac{1}{t}\left[(\phi_{-t}^X)_* Y(\phi_t^X(x)) - Y(x)\right]. \tag{5.6.3}$$

命题 5.6.5　在局部坐标下李导数(5.6.3) 可表示为

$$\mathrm{ad}_X(Y) = J_Y X - J_X Y = [X, Y]. \tag{5.6.4}$$

这里J_Y 是Y 的Jacobi 矩阵, 即

$$J_Y = \begin{bmatrix} \frac{\partial Y_1}{\partial x_1} & \cdots & \frac{\partial Y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial Y_n}{\partial x_1} & \cdots & \frac{\partial Y_n}{\partial x_n} \end{bmatrix}.$$

证明　根据Taylor 展式, 我们有

$$\phi_t^X(x) = x + (tX) + O(t^2). \tag{5.6.5}$$

$$Y(\phi_t^X(x)) = Y(x) + J_Y(tX) + O(t^2). \tag{5.6.6}$$

利用(5.6.5), ϕ_{-t}^X 的Jacobi 矩阵为

$$J_{\phi_{-t}^X} = I - tJ_X + O(t^2). \tag{5.6.7}$$

根据(5.6.5)~(5.6.7) 可得

$$
\begin{aligned}
(\phi^X_{-t})_* Y(\phi^X_t(x)) &= (I - tJ_X + O(t^2))(Y(x) + J_Y(tX) + O(t^2)) \\
&= Y(x) + t(J_Y X - J_X Y) + O(t^2).
\end{aligned}
$$

将其代入(5.6.3) 即得(5.6.4). □

定义 5.6.6 设$X \in V(M)$ 及$\alpha \in V^{*\omega}(M)$. 余向量场$\omega$ 关于X 的李导数, 记作$L_X(\alpha)$, 定义为

$$
L_X(\alpha) = \lim_{t\to 0} \frac{1}{t} \left[(\phi^X_t)^* \alpha(e^X_t(x)) - \omega(x) \right]. \tag{5.6.8}
$$

命题 5.6.7 在局部坐标下(5.6.5) 可表示为

$$
L_X(\alpha) = (J_{\alpha^{\mathrm{T}}} X)^{\mathrm{T}} + \alpha J_X. \tag{5.6.9}
$$

证明 类似于命题5.6.5 的证明, 我们先利用Taylor 展式得到

$$
\begin{aligned}
(\phi^X_t)^* \alpha(\phi^X_t(x)) &= (\alpha(x) + t(J_{\alpha^{\mathrm{T}}} X)^{\mathrm{T}} + O(t^2))(I + tJ_x + O(t^2)) \\
&= \alpha(x) + t(J_{\alpha^{\mathrm{T}}} X)^{\mathrm{T}} + t\alpha(x) J_X + O(t^2),
\end{aligned}
$$

这里转置来自以下的约定: 在局部坐标下余向量场总是表示成一个行向量. 将上式代入(5.6.8) 即得等式(5.6.9). □

高阶李导数可以用以下的递归方法来定义.

$$
L^{k+1}_X h = L^k_X(L_X h), \quad k \geq 1; \tag{5.6.10}
$$

$$
\mathrm{ad}^{k+1}_X Y = \mathrm{ad}^k_X(\mathrm{ad}_X Y), \quad k \geq 1; \tag{5.6.11}
$$

$$
L^{k+1}_X \alpha = L^k_X(L_X \alpha), \quad k \geq 1. \tag{5.6.12}
$$

下面考虑李导数的计算. 将函数$h \in C^\omega(M)$, 向量场$X, Y \in V^\omega(M)$ 及余向量场$\alpha \in V^{*\omega}(M)$ 作Taylor 展开, 我们有

$$
h = h_0 + h_1 x + h_2 x^2 + \cdots;
$$

$$
X = X_0 + X_1 x + X_2 x^2 + \cdots;
$$

$$
Y = Y_0 + Y_1 x + Y_2 x^2 + \cdots;
$$

$$
\alpha^{\mathrm{T}} = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots.
$$

为给出相应的公式, 我们需要以下引理, 这个引理本身也很有用. 直接计算就可以证明这个引理, 我们把它留给读者.

引理 5.6.8 设$X \in \mathbb{R}^n$ 为一列向量. 那么

$$X^{\mathrm{T}} = V_c^{\mathrm{T}}(I_n)X; \tag{5.6.13}$$

$$X = X^{\mathrm{T}}V_c^{\mathrm{T}}(I_n). \tag{5.6.14}$$

现在, 我们可以来推导李导数的Taylor 展式了.

命题 5.6.9 (1)

$$L_X h = \sum_{i=0}^{\infty} c_i x^i, \tag{5.6.15}$$

这里

$$c_i = \sum_{k=0}^{i} h_{k+1}\Phi_k^n(I_{n^k} \otimes X_{i-k}).$$

(2)

$$L_X Y = \sum_{i=0}^{\infty} d_i x^i, \tag{5.6.16}$$

这里

$$d_i = \sum_{k=0}^{i} \left[Y_{k+1}\Phi_k^n(I_{n^k} \otimes X_{i-k}) - X_{k+1}\Phi_k^n(I_{n^k} \otimes Y_{i-k})\right].$$

(3)

$$(L_X \alpha)^{\mathrm{T}} = \sum_{i=0}^{\infty} e_i x^i, \tag{5.6.17}$$

这里

$$e_i = \sum_{k=0}^{i} \left[\alpha_{k+1}\Phi_k^n(I_{n^k} \otimes X_{i-k}) - V_c^{\mathrm{T}}\left(I_{n^k} \otimes X_{k+1}\Phi_k^n\right)^{\mathrm{T}}\left(I_{n^k} \otimes \alpha_{i-k}\right)\right].$$

证明 我们只证明(5.6.17). 其余两式证明类似. 根据公式(5.6.9), 我们可得

$$(L_X \alpha)^{\mathrm{T}} = \frac{\partial \alpha^{\mathrm{T}}}{\partial x} X + \left(\frac{\partial X}{\partial x}\right)^{\mathrm{T}} \alpha^{\mathrm{T}}. \tag{5.6.18}$$

考虑其第一项:

$$\begin{aligned} \tfrac{\partial \alpha^{\mathrm{T}}}{\partial x} X &= \left(\sum_{i=1}^{\infty} \alpha_i \Phi_{i-1}^n x^{i-1}\right)\left(\sum_{i=0}^{\infty} X_i x^i\right) \\ &= \sum_{k=0}^{\infty}\left[\sum_{j=0}^{k} \alpha_{j+1}\Phi_j^n(I_{n^j} \otimes X_{k-j})\right] x^k. \end{aligned} \tag{5.6.19}$$

$$\begin{aligned}(\tfrac{\partial X}{\partial x})^{\mathrm{T}} &= \sum_{i=1}^{\infty}(x^{i-1})^{\mathrm{T}}(\Phi_{i-1}^{n})^{\mathrm{T}}X_i^{\mathrm{T}} \\ &= V_c^{\mathrm{T}}(I_{n^{i-1}})x^{i-1}(\Phi_{i-1}^{n})^{\mathrm{T}}X_i^{\mathrm{T}} \\ &= V_c^{\mathrm{T}}(I_{n^{i-1}})\left[I_{n^{i-1}}\otimes X_i\Phi_{i-1}^{n}\right]^{\mathrm{T}}x^{i-1}.\end{aligned}$$

因此

$$\begin{aligned}(\tfrac{\partial X}{\partial x})^{\mathrm{T}}\alpha^{\mathrm{T}} &= \sum_{i=1}^{\infty}\sum_{k=0}^{i}\left[V_c^{\mathrm{T}}I_{n^k}\left(I_{n^k}\otimes X_{k+1}\Phi_k^n\right)^{\mathrm{T}}x^k\alpha_{i-k}x^{i-k}\right] \\ &= \sum_{i=1}^{\infty}\left[\sum_{k=0}^{i}V_c^{\mathrm{T}}I_{n^k}\left(I_{n^k}\otimes X_{k+1}\Phi_k^n\right)^{\mathrm{T}}\left(I_{n^k}\otimes\alpha_{i-k}\right)\right]x^i.\end{aligned} \tag{5.6.20}$$

将(5.6.19) 及(5.6.20) 代入(5.6.18) 可得(5.6.17). □

5.7 注释与参考

多项式的半张量积表示, 实际上与协变张量的半张量积表示本质上是一致的. 它是用半张量积的工具, 或者更一般地说, 用矩阵方法解决非线性问题的关键.

利用多项式的半张量积表示, 多元函数的Taylor 展式也可以用张量积形式表示[1]. 半张量积形式表示的优势在于其结合律, 这可带来运算和讨论上的巨大方便.

关于经典Hamilton 系统、Hamilton 函数及Poisson 括号等概念, 在许多经典力学或物理学书中均有介绍, 例如文献[1].

目前, 控制理论研究最多的是伪Hamilton 系统与伪Hamilton 控制系统. 它与经典Hamilton 系统的最大区别是, 其结构矩阵很自由, 不必满足经典系统所要求的反对称性和Jacobi 恒等式等约束. 参见文献[98], [16].

李导数是非线性系统的几何理论研究中的一个基本工具, 不管对函数、向量场或余向量场的李导数均保持了导数的基本性质. 参见文献[68].

习 题 五

5.1 给出矩阵Ψ_3^3, 并用此检验例5.1.16 中$(A_3+DB_2)\Psi_3^3=0$ 成立.

5.2 给出$\nabla^k M(x)$ 与$D^k M(x)$ 之间的(类似(5.2.1) 的)块转置表示.

5.3 设$x=(x_1,\cdots,x_n)^{\mathrm{T}}\in\mathbb{R}^n$, 计算(a) Dx^2; (b) $D(Fx+Gx^2+Hx^3)$; (c) $\nabla(Fx+Gx^2+Hx^3)$.

5.4 给出公式$T_N(3,3)$ 与$T_B(3,3)$.

5.5 设

$$\begin{cases} y_1 = x_1 + x_2^2, \\ y_2 = \sin(x_1 + x_2). \end{cases}$$

给出逆映射的3次逼近, 即找出A_1, A_2, A_3, 使得

$$x = A_1 y + A_2 y^2 + A_3 y^3 + O(\|y\|^4).$$

5.6 设$h(x) \in C^\omega(N)$, $X, Y \in V^\omega(M)$, $\alpha \in V^{*\omega}(N)$. 利用它们的Taylor 展式, 给出2 阶李导数$L_X^2 h(x)$, $\mathrm{ad}_X^2 Y$ 及$L_X^2 \alpha$ 的Taylor 展式的系数公式.

第 6 章　逻辑的矩阵表示

数理逻辑是一门横跨哲学与自然科学的学科. 从自然科学的角度看, 它的产生是数学家们为揭示数学中的一些本质问题而努力的结果.

本章的目的是给出逻辑和多值逻辑的矩阵表示和逻辑运算的半张量积方法. 首先, 介绍逻辑的矩阵表示, 通过这种形式及半张量积运算可以证明逻辑的许多基本性质. 然后, 将这些结果推广到多值逻辑及其运算中去. 最后, 我们考虑利用这种方法来解决基于逻辑的智能系统中的问题和模糊逻辑控制的问题.

6.1　逻辑和它的矩阵表示

一个逻辑变量是指一个命题, 通常如果这个命题是真的, 我们就说该逻辑变量取值为“真” 或“1”, 如果这个命题是假的, 我们就说该逻辑变量取值为“假” 或“0”.

例 6.1.1　给定如下命题:

- A: 狗是四条腿的;
- B: 雪是黑的;
- C: 外星球存在高智商的生命.

显然, $A=1$, $B=0$, C 或者是1, 或者是0, 但只能是其中的一个, 只是我们现在还不知道它的值. □

在经典逻辑中, 逻辑变量只能取$\{0,1\}$ 这两个值. 但现实生活中,有时命题很难仅用“真” 或“假” 这两个值来刻画. 例如说: “某人是老人”. 如果此人是20 岁, 则显然命题是“假”, 如果此人是80 岁, 则显然命题是“真”. 但如果此人是40 岁或60 岁呢? 这时, 经典逻辑就显得无能为力了, 而模糊逻辑则可以给出回答. 模糊逻辑允许逻辑变量取$[0,1]$ 间的任何值, 因此, 如果此人是40 岁, 那么, 上述命题, 即逻辑变量, 可以取一个较小的值, 譬如, 0.2, 它表示命题不太对, 但不像20 岁那样完全不对. 如果此人是60 岁, 那么,逻辑变量可以取一个较大的值, 譬如, 0.8, 它表示命题比较合理, 但不像80 岁那样完全正确. 当然, 如何取值依赖于对隶属函数的选择, 这将在第7 节详述.

因此, 我们先考虑各种逻辑变量的取值空间.

定义 6.1.2　(1) (经典) 逻辑论域 D_ℓ 是指

$$D_\ell=\{T=1,F=0\};\tag{6.1.1}$$

(2) (经典) 逻辑变量P 是指在D_ℓ 中取值的变量, 即$P\in D_\ell$;

(3) 模糊逻辑论域 D_f 是指

$$D_f = \{\mu \,|\, 0 \leq \mu \leq 1\}; \tag{6.1.2}$$

(4) 模糊逻辑变量P 是指在D_f 中取值的变量, 即$P \in D_f$.

(5) k 值逻辑论域D_k 是指

$$D_k = \left\{ \frac{i}{k-1} \,\middle|\, i = 0, 1, \cdots, k-1 \right\}; \tag{6.1.3}$$

(6) k 值逻辑变量P 是指在D_k 中取值的变量, 即$P \in D_k$.

显然(经典) 逻辑变量是k 值逻辑变量在$k = 2$ 时的特例, 而k 值逻辑变量则是一种特殊的模糊逻辑变量.

我们记 $T_0 \equiv T \equiv 1$ 和 $F_0 \equiv F \equiv 0$ 分别表示"真"和"假".

定义 6.1.3 (1) r 元逻辑算子是指映射 $\sigma : \underbrace{D_f \times D_f \times \cdots \times D_f}_{r} \to D_f$.

(2) r 元k 值逻辑算子是指映射$\sigma : \underbrace{D_k \times D_k \times \cdots \times D_k}_{r} \to D_k$. 这里允许$k = 2$, 即经典逻辑的情况.

在本节的剩余部分我们仅考虑经典逻辑. 为了使用矩阵表示, 我们记

$$T := 1 \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad F := 0 \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{6.1.4}$$

定义 6.1.4 称2×2^r 矩阵M_σ 为r 元逻辑算子σ的结构矩阵, 如果

$$\sigma(P_1, \cdots, P_r) = M_\sigma \ltimes P_1 \ltimes \cdots \ltimes P_r := M_\sigma P_1 \cdots P_r. \tag{6.1.5}$$

注 本章所有的矩阵乘法都是半张量积, 因此后面我们将省略乘法符号$\ltimes$ (如(6.1.5) 中的最后一项).

首先, 我们考虑如何构造一个逻辑算子的结构矩阵. 考虑4 个基本的二元逻辑算子[87]: 析取, $P \vee Q$; 合取, $P \wedge Q$; 蕴涵, $P \to Q$; 等值, $P \leftrightarrow Q$.

它们的真值表如表6.1.1 所示[87].

根据真值表我们可以很容易得到这几个二元逻辑算子的结构矩阵. 以"$\vee$" (析取) 为例, 它在真值表中的值为$(1, 1, 1, 0)^{\mathrm{T}}$. 于是我们可以定义矩阵

$$M_d = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

表 6.1.1 真值表

P	Q	$P\vee Q$	$P\wedge Q$	$P\to Q$	$P\leftrightarrow Q$
1	1	1	1	1	1
1	0	1	0	0	0
0	1	1	0	1	0
0	0	0	0	1	1

通常, 如果r 元算子σ 的真值表是$(s_1,s_2,\cdots,s_{2^r})^{\mathrm{T}}$, 那么它的结构矩阵是:

$$M_\sigma=\begin{bmatrix} s_1 & s_2 & \cdots & s_{2^r}\\ 1-s_1 & 1-s_2 & \cdots & 1-s_{2^r}\end{bmatrix}. \tag{6.1.6}$$

下面的结论是很基本的, 它是真值表结构形式和半张量积结合律的一个直接结果.

定理 6.1.5 (6.1.6) 中定义的M_σ 是σ 的结构矩阵, 即对于M_σ, 等式(6.1.5) 成立.

证明 将M_σ 分成两个大小为$2\times 2^{r-1}$ 的块, 即$M_\sigma=(M_\sigma^1,M_\sigma^2)$. 于是(6.1.5) 的右边变成

$$\begin{bmatrix}M_\sigma^1 & M_\sigma^2\end{bmatrix}P_1P_2\cdots P_r=\begin{bmatrix}\left(M_\sigma^1 & M_\sigma^2P_1\right)\end{bmatrix}P_2\cdots P_r. \tag{6.1.7}$$

如果$P_1=(1,0)^{\mathrm{T}}$, 则(6.1.7) 变成

$$M_\sigma^1P_2\cdots P_r,$$

它对应于第一个变量$P_1=1$ 的块. 当$P_1=(0,1)^{\mathrm{T}}$ 时, (6.1.7) 变成

$$M_\sigma^2P_2\cdots P_r,$$

它对应于第一个变量$P_1=0$ 的块. 继续这个过程, 最后, 对于这些变量的相应取值, 可以得到在真值表中的精确值. □

利用定理6.1.5, 得到4 个基本二元逻辑算子的结构矩阵如下:

$$\begin{aligned} &M_\vee:=M_d=\begin{bmatrix}1&1&1&0\\0&0&0&1\end{bmatrix};\quad M_\wedge:=M_c=\begin{bmatrix}1&0&0&0\\0&1&1&1\end{bmatrix};\\ &M_\to:=M_i=\begin{bmatrix}1&0&1&1\\0&1&0&0\end{bmatrix};\quad M_\leftrightarrow:=M_e=\begin{bmatrix}1&0&0&1\\0&1&1&0\end{bmatrix}.\end{aligned} \tag{6.1.8}$$

下面我们讨论逻辑算子与它的结构矩阵之间的关系.

定理 6.1.6 (1) 一个逻辑算子 σ 具有唯一的结构矩阵.

(2) 一个2×2^r 矩阵是一个逻辑算子的结构矩阵, 当且仅当它的所有的列都属于$\mathcal{D}_\ell$.

证明 (1) 根据结构矩阵的构造, 由于对任一个σ, 它的真值表唯一, 于是它的结构矩阵也唯一.

(2) 设2×2^r 矩阵M_σ 为r 元逻辑算子σ 的结构矩阵, 即$\sigma(P_1, \cdots, P_r) = M_\sigma P_1 \cdots P_r$, 记$\xi_s$ 为M_σ 的第s 列. 将$s-1$ 表示成二进制形式$s_1 s_2 \cdots s_r$. (注意到由于$s-1 \leq 2^{2^r}-1$, 于是$s-1$ 的二进制的长度l 小于或等于2^r. 如果$l < 2^r$, 在它前面添加$2^r - l$ 个零以形成一个长度为2^r 的二进制数. 例如, 如果$r=2$ 并且$s=6$, 于是$s-1=101$, 那么$s_1 s_2 s_3 s_4 = 0101$.) 如果$s_i = 0$, 选择$P_i = (1,0)^{\mathrm{T}}$, 如果$s_i = 1$, 选择$P_i = (0,1)^{\mathrm{T}}$. (对于我们刚才$r=2$ 和$s=6$ 的例子有$P_1 = (1,0)^{\mathrm{T}}$, $P_2 = (0,1)^{\mathrm{T}}$, $P_3 = (1,0)^{\mathrm{T}}$, $P_4 = (0,1)^{\mathrm{T}}$.) 于是可以很容易检验

$$M_\sigma P_1 P_2 \cdots P_k = \xi_s,$$

因此$\xi_s \in \mathcal{D}_\ell$.

反之, 假设M 的所有列都属于$\mathcal{D}_\ell$, 并且记M 的第一行为M_1. 设σ 的真值表为M_1^{T}, 容易证明, 一个多线性映射的值由它在定义域即乘积空间的各因子空间的基底上的值唯一确定, 因此, 显然有$M = M_\sigma$. □

6.2 逻辑算子的一般结构

定理6.1.6 给出了一个逻辑算子集合的一般框架. 实际上, 对r 元算子我们知道有2^r 个不同情形的自变量值组合, 相应于每个自变量值组合一个算子又可能有2 种取值. 因此, 共有2^{2^r} 个不同的r 元逻辑算子. 一般地说, 对k 值逻辑, 有k^{k^r} 个不同的r 元k 值逻辑算子.

本节我们将对$r=1$ 和$r=2$ 的情况逐一考虑.

首先, 考虑$r=1$. 一般地说, 我们有4 个逻辑算子, 见表6.2.1.

表 6.2.1　一元逻辑算子

P	σ_0^1	σ_1^1	σ_2^1	σ_3^1
	F_0	$\neg$	$\equiv$	T_0
1	0	0	1	1
0	0	1	0	1

这里“F_0”为恒“假”算子, “T_0”为恒“真”算子, “$\neg$”为非(negation) 算子, “$\equiv$”为恒等(identity) 算子.

这些一元逻辑算子的结构矩阵分别是

$$M_{F_0}=\begin{bmatrix}0&0\\1&1\end{bmatrix};\quad M_n=\begin{bmatrix}0&1\\1&0\end{bmatrix};\quad M_{id}=\begin{bmatrix}1&0\\0&1\end{bmatrix};\quad M_{T_0}=\begin{bmatrix}1&1\\0&0\end{bmatrix}. \tag{6.2.1}$$

只有"¬" 经常使用, 但是为了完整性和后面的使用, 我们也对其他三种在此作了规定.

然后, 我们考虑二元逻辑算子: 除了在前面已经提过的4 种, 下面几种也是我们所熟知的:

- EOR (异或), $P\veebar Q$, 即当P 和Q 不同时取真时, 它取真, 否则取假;
- NAND (与非), $P\uparrow Q$, 即$P\uparrow Q\Leftrightarrow\neg(P\wedge Q)$;
- NOR (或非), $P\downarrow Q$, 即$P\downarrow Q\Leftrightarrow\neg(P\vee Q)$.

除了这7 种常见的二元逻辑算子外还有$2^{2^2}-7=9$ 种其他的二元逻辑算子. 我们将所有16 种二元逻辑算子总结在表6.2.2 中.

表 6.2.2 二元逻辑算子的真值表

P	Q	σ_0^2	σ_1^2	σ_2^2	σ_3^2	σ_4^2	σ_5^2	σ_6^2	σ_7^2
		F_0	$\downarrow$	$-^*$	$\neg_1$	$-$	$\neg_2$	$\veebar$	$\uparrow$
1	1	0	0	0	0	0	0	0	0
1	0	0	0	0	0	1	1	1	1
0	1	0	0	1	1	0	0	1	1
0	0	0	1	0	1	0	1	0	1
P	Q	σ_8^2	σ_9^2	σ_{10}^2	σ_{11}^2	σ_{12}^2	σ_{13}^2	σ_{14}^2	σ_{15}^2
		$\wedge$	$\leftrightarrow$	Q	$\rightarrow$	P	$\rightarrow^*$	$\vee$	T_0
1	1	1	1	1	1	1	1	1	1
1	0	0	0	0	0	1	1	1	1
0	1	0	0	1	1	0	0	1	1
0	0	0	1	0	1	0	1	0	1

其中$\sigma_0^2=F_0$, $\sigma_{15}^2=T_0$ 是常值逻辑算子; $\sigma_3^2=\neg P$, $\sigma_5^2=\neg Q$, $\sigma_{10}^2=Q$ 和$\sigma_{12}^2=P$ 就是一元逻辑算子. 我们有下面的新的二元逻辑算子:

- EOR (σ_6^2, 异或) ($P\veebar Q$):

$$P\veebar Q\Leftrightarrow\neg(P\leftrightarrow Q); \tag{6.2.2}$$

- NAND (σ_7^2, 与非) ($P\uparrow Q$):

$$P\uparrow Q\Leftrightarrow\neg(P\wedge Q); \tag{6.2.3}$$

- NOR (σ_1^2, 或非) ($P \downarrow Q$):

$$P \downarrow Q \Leftrightarrow \neg(P \vee Q); \tag{6.2.4}$$

- NIMP (σ_4^2, 非蕴涵) ($P - Q$):

$$P - Q \Leftrightarrow \neg(P \to Q); \tag{6.2.5}$$

- NIIMP (σ_2^2, 非逆蕴涵) ($P -^* Q$):

$$P -^* Q \Leftrightarrow \neg(Q \to P); \tag{6.2.6}$$

- ICOD (σ_{13}^2, 逆蕴涵) $P \to^* Q$:

$$P \to^* Q \Leftrightarrow Q \to P. \tag{6.2.7}$$

(6.2.7) 的等价表示可以直接由定义得到, 而(6.2.2)~(6.2.6) 的等价我们将在后面证明.

另外, 为了对这些算子使用一些有意义的符号, 我们引入记号σ_j^i 以便于我们使用. 这里上标i 表示算子的阶数, 而将下标j 表示序数, 即第几个i 阶算子. 有趣的是, 当j 转换成二进制就是算子的真值表. 利用这个记号就可以直接得到σ_j^i 的结构矩阵.

例 6.2.1　考虑$\neg_2$, 它是第二个变量的否定式, 它的另一个记号是σ_5^2. 由于$5 = 101 = 0101$, 因此

$$M_{\neg_2} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

□

6.3　基本逻辑算子的性质

根据基本逻辑算子的真值表的结构, 我们有下面的否定性质.

命题 6.3.1　给定一个r 元逻辑算子σ_a^r, 它的否定算子是$\sigma_{2^{2^r}-a-1}^r$, 即

$$\neg\sigma_a^r(P,Q) = \sigma_{2^{2^r}-a-1}^r(P,Q). \tag{6.3.1}$$

证明　由于$2^{2^r} - 1 = \underbrace{1\,1\,\cdots\,1}_{2^r}$. 将$a$ 表示成二进制形式(可能需要在前面添加一些零) 的2^r 个数字的二进制数为

$$a = a_1 a_2 \cdots a_{2^r}.$$

于是$(a_1, a_2, \cdots, a_{2^r})^{\mathrm{T}}$ 是σ_a^r 的真值表. 用二进制形式有$2^{2^r}-1-a=[(1-a_1),(1-a_2),\cdots,(1-a_r)]$. 另一方面, $\sigma^r_{2^{2^r}-a-1}$ 的真值表是$(1-a_1, 1-a_2, \cdots, 1-a_r)^{\mathrm{T}}$. 因此(6.3.1) 成立. □

注 利用命题6.3.1, 我们立即有(6.3.2)~(6.3.6).

定义 6.3.2 (1) 称两个逻辑表达式是逻辑等价的, 如果当变量在$\mathcal{D}_\ell$ 中取任意值时它们的值都相等.

(2) 称两个逻辑表达式是绝对逻辑等价的, 如果当变量在$\mathcal{D}_f$ 中取任意值时它们的值都相等.

设两个逻辑表达式L_1 与L_2 逻辑等价, 我们记为$L_1 \Leftrightarrow L_2$. 注意"逻辑等价" (tautological equivalence) 与"等值" (bicondition, $L_1 \leftrightarrow L_2$) 是完全不同的概念.

命题 6.3.3 设两个逻辑表达式含有相同的逻辑变量, 并且每一个变量仅出现一次. 那么它们绝对逻辑等价, 当且仅当它们逻辑等价.

证明 可以看出, 在矩阵表示下, 一个逻辑表达式关于每个变量都是一个线性映射, 只要这个变量只在表达式中仅出现一次. 现在这个条件可以保证两个表达式关于它的所有变量是多线性的. 如果对于$P_i=(1,0)^{\mathrm{T}}$ 和$P_i=(0,1)^{\mathrm{T}}$, 它们的矩阵表示一样, 那么线性组合就表明对于$P_i=(\mu, 1-\mu)^{\mathrm{T}}$, $0 \le \mu \le 1$, 它们是逻辑等价的. □

命题 6.3.4 下面的式子都是绝对逻辑等价的:

$$\neg\neg P \Leftrightarrow P. \tag{6.3.2}$$

$$(P \wedge Q) \wedge R \Leftrightarrow P \wedge (Q \wedge R). \tag{6.3.3}$$

$$(P \vee Q) \vee R \Leftrightarrow P \vee (Q \vee R). \tag{6.3.4}$$

$$\neg(P \wedge Q) = \neg P \vee \neg Q. \tag{6.3.5}$$

$$\neg(P \vee Q) \Leftrightarrow \neg P \wedge \neg Q. \tag{6.3.6}$$

$$P \to Q \Leftrightarrow \neg P \vee Q. \tag{6.3.7}$$

$$\neg(P \to Q) \Leftrightarrow P \wedge \neg Q. \tag{6.3.8}$$

$$P \to Q \Leftrightarrow \neg Q \to \neg P. \tag{6.3.9}$$

$$P \to (Q \to R) \Leftrightarrow (P \wedge Q) \to R. \tag{6.3.10}$$

$$\neg(P \leftrightarrow Q) \Leftrightarrow P \leftrightarrow \neg Q. \tag{6.3.11}$$

证明　我们只证明(6.3.9), 其余类似. 由命题6.3.3, 我们只需要证明它是逻辑等价的即可.

$$\begin{aligned} RHS &= M_iM_nQM_nP = M_iM_n(I_2 \otimes M_n)QP \\ &= M_iM_n(I_2 \otimes M_n)W_{[2]}PQ. \end{aligned}$$

由于

$$\begin{aligned} &M_iM_n(I_2 \otimes M_n)W_{[2]} \\ &= \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} = M_i, \end{aligned}$$

于是(6.3.9) 成立. □

定义 6.3.5　称一个r 元算子是对称的, 如果

$$M_\sigma P_1P_2\cdots P_k = M_\sigma P_{\lambda(1)}P_{\lambda(2)}\cdots P_{\lambda(k)}, \quad \forall \lambda \in \mathbf{S}_k, \tag{6.3.12}$$

其中$\mathbf{S}_k$ 是k 阶置换群.

命题 6.3.6　一个二元基本表达式是对称的, 当且仅当在它的真值表$(s_1, s_2, s_3, s_4)^{\mathrm{T}}$ 中

$$s_2 = s_3.$$

证明　注意到表达式σ 的结构矩阵是

$$M_\sigma = \begin{bmatrix} s_1 & s_2 & s_3 & s_4 \\ 1-s_1 & 1-s_2 & 1-s_3 & 1-s_4 \end{bmatrix}.$$

$$\begin{aligned} \sigma(P,Q) &= M_\sigma PQ = M_\sigma W_{[2]}QP \\ &= \begin{bmatrix} s_1 & s_3 & s_2 & s_4 \\ 1-s_1 & 1-s_3 & 1-s_2 & 1-s_4 \end{bmatrix} QP \\ &= M_\sigma QP = \sigma(Q,P). \end{aligned}$$

□

例 6.3.7 考查表6.2.2 中的二元算子. 直接计算可以得到F_0, ↓, ⊻, ↑, ∧, ↔, ∨ 和T_0 是对称的. □

命题 6.3.8 下列式子是逻辑等价的:

$$P \vee P \Leftrightarrow P. \tag{6.3.13}$$

$$P \wedge P \Leftrightarrow P. \tag{6.3.14}$$

$$R \vee (P \wedge \neg P) \Leftrightarrow R. \tag{6.3.15}$$

$$R \wedge (P \vee \neg P) \Leftrightarrow R. \tag{6.3.16}$$

$$P \wedge (Q \vee R) \Leftrightarrow (P \wedge Q) \vee (P \wedge R). \tag{6.3.17}$$

$$P \vee (Q \wedge R) \Leftrightarrow (P \vee Q) \wedge (P \vee R). \tag{6.3.18}$$

$$P \leftrightarrow Q \Leftrightarrow (P \rightarrow Q) \wedge (Q \rightarrow P). \tag{6.3.19}$$

$$P \leftrightarrow Q \Leftrightarrow (P \wedge Q) \vee (\neg P \wedge \neg Q). \tag{6.3.20}$$

证明 我们只证明(6.3.15). 设$R = (\delta, 1-\delta)^{\mathrm{T}}$, $P = (\mu, 1-\mu)^{\mathrm{T}}$, 则

$$\begin{aligned}
LHS &= M_{\vee} R M_{\wedge} P M_{\neg} P \\
&= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta \\ 1-\delta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ 1-\mu \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \mu \\ 1-\mu \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta \\ 1-\delta \end{bmatrix} \begin{bmatrix} \mu(1-\mu) \\ \mu^2 + \mu(1-\mu) + (1-\mu)^2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \delta[\mu(1-\mu)] \\ \delta[\mu^2 + \mu(1-\mu) + (1-\mu)^2] \\ (1-\delta)[\mu(1-\mu)] \\ (1-\delta)[\mu^2 + \mu(1-\mu) + (1-\mu)^2] \end{bmatrix} \\
&= \begin{bmatrix} \delta + (1-\delta)\mu(1-\mu) \\ (1-\delta)[\mu^2 + \mu(1-\mu) + (1-\mu)^2] \end{bmatrix}.
\end{aligned}$$

只要$\mu \in \{0, 1\}$, 我们就有$LHS = (\delta, 1-\delta)^{\mathrm{T}} = R$. □

容易检验, (6.3.13)~(6.3.20) 都不是绝对逻辑等价的, 例如对(6.3.15) 在上述证明中如果取$\mu = 0.5$, 则$LHS = \begin{bmatrix} 0.25 + 0.75\delta \\ 0.75 - 0.75\delta \end{bmatrix} \neq R$.

定义 6.3.9 设两个逻辑表达式L_1, L_2 满足$L_1 \to L_2$ 恒真(即为重言式), 则称L_1 到L_2 的逻辑推断成立, 记作$L_1 \Rightarrow L_2$.

注 "逻辑推断" (tautological implication) 与 "蕴涵" (condition,$L_1 \to L_2$) 是完全不同的概念.

命题 6.3.10 $L_1 \Rightarrow L_2$, 当且仅当$L_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ 时, $L_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

证明 (充分性) 设$L_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $L_1 = \begin{bmatrix} \alpha \\ 1-\alpha \end{bmatrix}$, 则

$$L_1 \to L_2 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ 1-\alpha \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

如果$L_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, 依条件$L_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, 则

$$L_1 \to L_2 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

(必要性) 设$L_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, 但$L_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, 则

$$L_1 \to L_2 = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

□

注 上述命题的物理意义是清楚的, 即只要L_1 "真", 则必有L_2 "真".

在下面的命题证明中, 我们将利用命题6.3.10.

命题 6.3.11 下面的式子逻辑推断成立.

$$P \wedge Q \Rightarrow P. \tag{6.3.21}$$

$$P \wedge Q \Rightarrow Q. \tag{6.3.22}$$

$$P \Rightarrow P \vee Q. \tag{6.3.23}$$

$$Q \Rightarrow P \vee Q. \tag{6.3.24}$$

$$\neg P \Rightarrow P \rightarrow Q. \tag{6.3.25}$$

$$Q \Rightarrow P \rightarrow Q. \tag{6.3.26}$$

$$\neg(P \rightarrow Q) \Rightarrow P. \tag{6.3.27}$$

$$\neg(P \rightarrow Q) \Rightarrow \neg Q. \tag{6.3.28}$$

$$\neg P \wedge (P \vee Q) \Rightarrow Q. \tag{6.3.29}$$

$$P \wedge (P \rightarrow Q) \Rightarrow Q. \tag{6.3.30}$$

$$\neg Q \wedge (P \rightarrow Q) \Rightarrow \neg P. \tag{6.3.31}$$

$$(P \rightarrow Q) \wedge (Q \rightarrow R) \Rightarrow P \rightarrow R. \tag{6.3.32}$$

$$(P \vee Q) \wedge (P \rightarrow R) \wedge (Q \rightarrow R) \Rightarrow R. \tag{6.3.33}$$

证明 我们只证明(6.3.33), 利用命题6.3.10, 假设右边假, 即$R = (0,1)^{\mathrm{T}}$, 我们检验左边:

$$\begin{aligned}
& (P \vee Q) \wedge (P \rightarrow R) \wedge (Q \rightarrow R) \\
= \; & M_{\wedge} M_{\vee} P Q M_{\wedge} M_{\rightarrow} P R M_{\rightarrow} Q R \\
= \; & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p \\ 1-p \end{bmatrix} \begin{bmatrix} q \\ 1-q \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \\
& \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} p \\ 1-p \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} q \\ 1-q \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\
= \; & \begin{bmatrix} (p+q-pq)(1-p)(1-q) \\ (p+q-pq)^2 + (1-p)^2(1-q)^2 + (p+q-pq)(1-p)(1-q) \end{bmatrix}.
\end{aligned}$$

我们需要检验4种情形: (1) $p=0, q=0$, (2) $p=0, q=1$, (3) $p=1, q=0$, (4) $p=1, q=1$. 不论哪种情形最后一个矩阵都是$(0,1)^{\mathrm{T}}$. □

最后, 我们利用矩阵表示考虑一些关于EOR、NAND 和NOR 的等价. 下面式子的证明都是类似的.

命题 6.3.12　下面的式子都是逻辑等价的:

$$P \veebar Q \Leftrightarrow Q \veebar P. \tag{6.3.34}$$

$$(P \veebar Q) \veebar R \Leftrightarrow P \veebar (Q \veebar R). \tag{6.3.35}$$

$$P \wedge (Q \veebar R) \Leftrightarrow (P \wedge Q) \veebar (P \wedge R). \tag{6.3.36}$$

$$P \veebar Q \Leftrightarrow (P \wedge \neg Q) \vee (\neg P \wedge Q). \tag{6.3.37}$$

$$P \veebar Q \Leftrightarrow \neg(P \leftrightarrow Q). \tag{6.3.38}$$

$$P \uparrow Q \Leftrightarrow Q \uparrow P. \tag{6.3.39}$$

$$P \downarrow Q \Leftrightarrow Q \downarrow P. \tag{6.3.40}$$

$$P \uparrow (Q \uparrow R) \Leftrightarrow \neg P \vee (Q \wedge R). \tag{6.3.41}$$

$$(P \uparrow Q) \uparrow R \Leftrightarrow (P \wedge Q) \vee \neg R. \tag{6.3.42}$$

$$P \downarrow (Q \downarrow R) \Leftrightarrow \neg P \wedge (Q \vee R). \tag{6.3.43}$$

$$(P \downarrow Q) \downarrow R \Leftrightarrow (P \vee Q) \wedge \neg R. \tag{6.3.44}$$

6.4　逻辑表达式的规范型

本节我们将继续考虑命题演算, 并给出命题的逻辑表达式的规范型.

引理 6.4.1 一个含有r 个逻辑变量的逻辑命题

$$\sigma(A_1, \cdots, A_r) \tag{6.4.1}$$

总可以表示成

$$\sigma(A_1, \cdots, A_k) = M_\sigma A_1^{p_1} A_2^{p_2} \cdots A_r^{p_r}, \tag{6.4.2}$$

其中 M_σ 是一个 $2 \times 2^{p_1+\cdots+p_r}$ 结构矩阵.

证明 利用每个逻辑算子的矩阵表示, 一个逻辑表达式总可以表示成

$$\sigma(A_1, \cdots, A_k) = M_{i_1} A_{i_1} \cdots M_{i_s} A_{i_s}, \tag{6.4.3}$$

其中

$$A_{i_j} \in \{A_1, A_2, \cdots, A_k\}.$$

既然矩阵的半张量积具有结合律, 利用换位公式(3.3.21), (6.4.3) 可以改写成

$$\sigma(A_1, \cdots, A_k) = M A_{i_1} \cdots A_{i_s}. \tag{6.4.4}$$

然后利用换位矩阵的性质, 由命题2.3.19 可知: 设$A_i \in \mathbb{R}^n$, $i = 1, \cdots, s$. 定义

$$P = \underbrace{I_n \otimes \cdots \otimes I_n}_{i-1} \otimes W_{[n]} \otimes \underbrace{I_n \otimes \cdots \otimes I_n}_{s-i-1}.$$

于是

$$P A_1 \cdots A_i A_{i+1} \cdots A_s = A_1 \cdots A_{i+1} A_i \cdots A_s. \tag{6.4.5}$$

注意到$P^{-1} = P$, 再利用(6.4.5), 很容易将(6.4.4) 转换成(6.4.2) 的形式. □

引理 6.4.2 设一个逻辑命题的矩阵表示为MA^2, 其中 M 是一个 $p \times 4q$ 矩阵. 将 M 分割为 $M = \begin{bmatrix} M_1 & M_2 & M_3 & M_4 \end{bmatrix}$, 其中 M_i 是 $p \times q$ 矩阵, 那么

$$MA^2 = \begin{bmatrix} M_1 & M_4 \end{bmatrix} A. \tag{6.4.6}$$

证明 利用矩阵半张量积的性质, 容易验证(6.4.6) 对于 $A = (1, 0)^{\mathrm{T}}$ 和 $A = (0, 1)^{\mathrm{T}}$ 都成立. □

由引理6.4.2, 我们定义降幂矩阵为

$$M_r = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{6.4.7}$$

它的名称来自于它下面的性质, 这个性质是引理6.4.2 的一个直接结果.

引理 6.4.3 设 A 是一个逻辑变量, 那么对于任意的 $p \times 4q$ 矩阵 Ψ 有

$$\Psi A^2 = \Psi M_r A. \tag{6.4.8}$$

在一个逻辑表达式中称一个逻辑变量是常量, 如果它的值是预先给定的. 称它为自由变量, 如果它的值是任意的. 利用这个概念和上面的引理, 我们有

定理 6.4.4 任一个含有自由变量$P_1, \cdots, P_r$的逻辑表达式$L(P_1, \cdots, P_r)$ 都可以唯一地表示成规范型

$$L(P_1, \cdots, P_r) = M_L P_1 P_2 \cdots P_r, \tag{6.4.9}$$

其中M_L 是一个2×2^r 结构矩阵.

证明 利用引理6.4.1, 将矩阵表示转换为(6.4.2). 然后利用引理6.4.3, 即(6.4.8) 式, 将每个变量的指数降为 1. 注意到半张量积可以自动满足维数的需要, 于是表达式(6.4.9) 是关于不同变量的一个多线性形式, 它一定是唯一的. □

作为一个应用, 我们用 M_c, M_d 和 M_n 来表示 M_i 和 M_e.

命题 6.4.5 M_i 和 M_e 可以用 M_c, M_d 和 M_n 分别表示成

$$M_i = M_d M_n; \tag{6.4.10}$$

和

$$\begin{aligned} M_e &= M_c M_i (I_4 \otimes M_i)(I_2 \otimes M_r)(I_2 \otimes W_{[2]}) M_r \\ &= M_c M_d M_n (I_4 \otimes M_d M_n)(I_2 \otimes M_r)(I_2 \otimes W_{[2]}) M_r. \end{aligned} \tag{6.4.11}$$

证明 注意到逻辑等价

$$P \to Q \Leftrightarrow \neg P \vee Q.$$

将它表示成矩阵形式, 我们有

$$M_i PQ = M_d (M_n P) Q = M_d M_n PQ.$$

于是(6.4.10) 成立. 至于(6.4.11), 我们有

$$P \leftrightarrow Q \Leftrightarrow (P \to Q) \wedge (Q \to P).$$

表示成矩阵形式

$$M_e PQ = M_c (M_i PQ)(M_i QP).$$

现在右边是

$$\begin{aligned} M_cM_iPQM_iQP &= M_cM_dM_n(I_4 \otimes M_dM_n)PQQP \\ &= M_cM_dM_n(I_4 \otimes M_dM_n)P(M_r)QP \\ &= M_cM_dM_n(I_4 \otimes M_dM_n)(I_2 \otimes M_r)PW_{[2]}PQ \\ &= M_cM_dM_n(I_4 \otimes M_dM_n)(I_2 \otimes M_r)(I_2 \otimes W_{[2]})PPQ \\ &= M_cM_dM_n(I_4 \otimes M_dM_n)(I_2 \otimes M_r)(I_2 \otimes W_{[2]})M_rPQ. \end{aligned}$$

(6.4.11)得证. □

注 利用(6.4.10), (6.4.11) 和逻辑等价关系式(6.2.2)~(6.2.7), 可以看出所有二元逻辑算子的矩阵都可以用M_c, M_d 和M_n 表示. 我们对此并不感到奇怪, 因为$\{\vee, \wedge, \neg\}$ 是一个完全集, 即其他逻辑算子都可由它们表示出来. 在后面, 我们还会用它们来定义多值逻辑或混合值逻辑下的相应算子.

下面我们用一个简单的例子来说明定理6.4.4 的应用.

例 6.4.6 甲说乙撒谎, 乙说丙撒谎, 而丙说甲和乙都撒谎, 那么到底谁在撒谎?

记A: 甲没有撒谎; B: 乙没有撒谎; C: 丙没有撒谎. 那么上面的陈述的逻辑表达式是

$$(A \leftrightarrow \neg B) \wedge (B \leftrightarrow \neg C) \wedge (C \leftrightarrow \neg A \wedge \neg B). \tag{6.4.12}$$

问题就变成何时(6.4.12) 是真. (6.4.12) 的矩阵表示是

$$\begin{aligned} & M_c^2 M_e A M_n B M_e B M_n C M_e C M_c M_n A M_n B \\ = & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}^2 \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} A \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} B \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} B \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} C \\ & \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} C \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} A \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} B. \end{aligned} \tag{6.4.13}$$

现在我们需要将它降幂为形如(6.4.9) 的规范型. 这个简化的过程可以交给计算机

来完成. 我们在这里简单地演示一下这个过程:

$$\begin{aligned} M_cM_nAM_nB &= M_cM_n(I_2 \otimes M_n)AB \\ &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} \left(I_2 \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) AB \\ &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} AB. \end{aligned}$$

$$\begin{aligned} M_nCM_eC &= M_n[I_2 \otimes M_e]C^2 = M_n[I_2 \otimes M_e]M_rC \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} C \\ &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} C \\ &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} C. \end{aligned}$$

同理,

$$M_nBM_eB = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} B.$$

现在(6.4.13) 变成

$$M_c^2M_eA \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} B \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} C \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} AB. \tag{6.4.14}$$

由于

$$C\begin{bmatrix}0&0&0&1\\1&1&1&0\end{bmatrix}AB=\left(I_2\otimes\begin{bmatrix}0&0&0&1\\1&1&1&0\end{bmatrix}\right)CAB.$$

(6.4.13) 变成

$$M_c^2M_eA\begin{bmatrix}0&0&0&1\\0&0&1&0\\1&0&0&0\\0&1&0&0\end{bmatrix}B\begin{bmatrix}0&0&0&0&1&1&1&0\\0&0&0&0&0&0&0&1\\0&0&0&1&0&0&0&0\\1&1&1&0&0&0&0&0\end{bmatrix}CAB. \tag{6.4.15}$$

利用换位矩阵, 我们有

$$CAB=W_{[2,4]}W_{[4,2]}CAB=W_{[2,4]}BCA.$$

现在

$$W_{[2,4]}=\begin{bmatrix}1&0&0&0&0&0&0&0\\0&0&0&0&1&0&0&0\\0&1&0&0&0&0&0&0\\0&0&0&0&0&1&0&0\\0&0&1&0&0&0&0&0\\0&0&0&0&0&0&1&0\\0&0&0&1&0&0&0&0\\0&0&0&0&0&0&0&1\end{bmatrix}.$$

将它们代入到(6.4.15), (6.4.13) 就变成

$$M_c^2M_eA\begin{bmatrix}0&0&1&1&0&0&0&0\\0&0&0&0&0&0&0&0\\0&0&0&0&0&0&0&0\\1&1&0&0&0&0&0&0\\0&0&0&0&0&0&1&0\\0&0&0&0&0&0&0&1\\0&0&0&0&0&1&0&0\\0&0&0&0&1&0&0&0\end{bmatrix}BCA. \tag{6.4.16}$$

对A 使用相同的技巧(先交换最后的A 和它前面的BC, 再交换第一个A 和它后面的常数矩阵, 使两个A 因子移到一起, 然后并将A^2 降幂成A), 再直接计算可得:

(6.4.13) 等于

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 1 \end{bmatrix} ABC. \tag{6.4.17}$$

(6.4.17) 就是(6.4.13) 的规范型. 为了使(6.4.17) 取真(即(6.4.17) = $(1,0)^{\mathrm{T}}$), 唯一的解是

$$A = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

因此, 甲和丙都撒谎了, 只有乙没有撒谎. □

下面的两个公式可以用来降低计算的复杂性, 利用命题2.3.19 很容易证明它们.

命题 6.4.7

$$W_{[n,n^k]} = \prod_{i=0}^{k-1} I_{n^{k-1-i}} \otimes W_{[n]} \otimes I_{2^i}; \tag{6.4.18}$$

$$W_{[n^k,n]} = \prod_{i=k-1}^{0} I_{n^{k-1-i}} \otimes W_{[n]} \otimes I_{2^i}. \tag{6.4.19}$$

6.5 多值逻辑

逻辑的矩阵形式的另一个优势是它可以很容易地推广到多值逻辑. 设P 和Q 是两个k 值逻辑变量, 即它们从$\mathcal{D}_k$ ($k \geq 2$) 中取值. 第一个推广是, 利用变量的逻辑值, 我们定义一些基本算子如下:

定义 6.5.1 设P 和Q 是两个k 值逻辑变量. 定义它们的析取为

$$P \vee Q = \max(P, Q); \tag{6.5.1}$$

定义它们的合取为

$$P \wedge Q = \min(P, Q). \tag{6.5.2}$$

定义否定为

$$\neg P = 1 - P. \tag{6.5.3}$$

为了使用矩阵表示, 我们需要将逻辑值表示成向量形式, 与(6.1.4) 类似, 我们定义

$$\frac{i}{k-1} \equiv \delta_{k-i}^{k}, \quad i = 0, 1, \cdots, k-1, \tag{6.5.4}$$

其中δ_j^k 是单位阵I_k 的第j 列. 确切地, 我们有

$$1 \equiv \begin{bmatrix}1 & 0 & \cdots & 0\end{bmatrix}^{\mathrm{T}};\ \frac{k-2}{k-1} \equiv \begin{bmatrix}0 & 1 & \cdots & 0\end{bmatrix}^{\mathrm{T}};\ \cdots;\ 0 \equiv \begin{bmatrix}0 & 0 & \cdots & 1\end{bmatrix}^{\mathrm{T}}.$$

利用逻辑值的向量形式, 可以很容易算出定义6.5.1 中定义的算子的结构矩阵.

为了记号的方便, 在本节的大多数地方我们假设$k=3$. 所有的结果都适用于$k \geq 2$ 的情形.

命题 6.5.2 假设$k=3$, 那么"析取"、"合取"和"否定"的结构矩阵分别是

$$M_d^3 = \begin{bmatrix}1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}. \tag{6.5.5}$$

$$M_c^3 = \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1\end{bmatrix}. \tag{6.5.6}$$

$$M_n^3 = \begin{bmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{bmatrix}. \tag{6.5.7}$$

定义6.5.1 是经典逻辑的一个自然推广和应用. 它在$k=2$ 时显然与经典逻辑的相应运算的定义一致, 但是蕴涵有多种形式的定义. 因此有很多种不同的3 值逻辑. 例如, 3 种常见的3 值逻辑Kleene-Dienes 型(KD), Lukasiewicz 型(L), Bochvar 型(B) 如表6.5.1 所示($T=1$, $U=0.5$, $F=0$)[79].

表 6.5.1 3 值逻辑

		KD		L		B	
P	Q	$\rightarrow$	$\leftrightarrow$	$\rightarrow$	$\leftrightarrow$	$\rightarrow$	$\leftrightarrow$
T	T	T	T	T	T	T	T
T	U	U	U	U	U	U	U
T	F	F	F	F	F	F	F
U	T	T	U	T	U	U	U
U	U	U	U	T	T	U	U
U	F	U	U	U	U	U	U
F	T	T	F	T	F	T	F
F	U	T	U	T	U	U	U
F	F	T	T	T	T	T	T

对于k 值逻辑, 同样可以通过结构矩阵来定义逻辑算子.

现在我们用(6.4.10) 来定义蕴涵. 当$k=3$ 时, 容易算出

$$M_i^3=\begin{bmatrix}1&0&0&1&0&0&1&1&1\\0&1&0&0&1&1&0&0&0\\0&0&1&0&0&0&0&0&0\end{bmatrix}. \tag{6.5.8}$$

我们还可以用(6.4.11) 来定义等值. 为此需要一个记号. 设δ_i^k 是I_k 的第i 列, $i=1,\cdots,k$. 易证k 值逻辑的降幂矩阵是

$$M_r^k=\begin{bmatrix}\delta_1^k&0&\cdots&0\\0&\delta_2^k&\cdots&0\\0&0&\ddots&0\\0&0&\cdots&\delta_s^k\end{bmatrix}. \tag{6.5.9}$$

注意到利用相同的论证可以看出, 对于k 值逻辑(6.4.11) 变成

$$M_e=M_cM_i(I_{k^2}\otimes M_i)(I_k\otimes M_r)(I_k\otimes W_{[k]})M_r. \tag{6.5.10}$$

利用这个公式, 我们可以算出

$$M_e^3=\begin{bmatrix}1&0&0&0&0&0&0&0&1\\0&1&0&1&1&1&0&1&0\\0&0&1&0&0&0&1&0&0\end{bmatrix}. \tag{6.5.11}$$

显然这个蕴涵所定义的3 值逻辑就是Kleene-Dienes 型逻辑.

我们给出下面的例子来说明如何应用矩阵表示来进行模糊逻辑推理.

例 6.5.3 [79]　一个侦探对于一个谋杀案有如下线索:

(1) 80% 可以确定A 或B 是凶手;

(2) 如果A 是凶手, 很可能案发时间不在午夜之前;

(3) 如果B 的供词正确, 那么房间的灯在午夜时仍亮着;

(4) 如果B 的供词不正确, 那么很可能案发时间在午夜之前;

(5) 有证据表明房间的灯在午夜时关着.

我们假设(作为通常的理解) “很可能”的可能性高于“80%”的可能性, 并且量化“T”, “很可能”, “80%”, “1% ~ 80%”, “很不可能”, “F” 为6 个逻辑层次, 于是我们可以在6 值逻辑上考虑这个问题. 记命题为

(1) A: A 是凶手;

(2) B: B 是凶手;

(3) M: 案发时间在午夜之前;

(4) S: B 的供词正确;

(5) L: 午夜时房间的等亮着.

于是我们有下面的模糊逻辑方程

$$\begin{aligned} A \vee B &= (0,0,1,0,0,0)^{\mathrm{T}}, \\ A \to \neg M &= (0,1,0,0,0,0)^{\mathrm{T}}, \\ S \to L &= (1,0,0,0,0,0)^{\mathrm{T}}, \\ \neg S \to M &= (0,1,0,0,0,0)^{\mathrm{T}}, \\ \neg L &= (1,0,0,0,0,0)^{\mathrm{T}}. \end{aligned} \tag{6.5.12}$$

我们利用矩阵表示来进行模糊逻辑推理. 首先, 由$\neg L = (1,0,0,0,0,0)^{\mathrm{T}}$ 我们有

$$L = (0,0,0,0,0,1)^{\mathrm{T}}.$$

再由$S \to L = (1,0,0,0,0,0)^{\mathrm{T}}$ 我们有如下矩阵形式

$$M_i^6 SL = M_i^6 W_{[6]} LS := \Psi_1 S = (1,0,0,0,0,0)^{\mathrm{T}};$$

容易算出

$$\Psi_1 = M_i^6 W_{[6]} L = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

于是可以解出$S = (0,0,0,0,0,1)^{\mathrm{T}}$. 同理, 由$\neg S \to M = (0,1,0,0,0,0)^{\mathrm{T}}$ 有

$$M_i^6 M_n^6 SM = (0\,1\,0\,0\,0\,0)^{\mathrm{T}}.$$

可以解出

$$M = (0,1,0,0,0,0)^{\mathrm{T}}.$$

考虑$A \to \neg M = M_i^6 A M_n^6 M = (0,1,0,0,0,0)^{\mathrm{T}}$. 利用半张量积的一些性质, 我们有

$$M_i^6 A M_n^6 M = M_i^6 (I_6 \otimes M_n^6) AM = M_i^6 (I_6 \otimes M_n^6) W_{[6]} MA := \psi_2 A.$$

容易算出

$$\psi_2 = M_i^6(I_6 \otimes M_n^6)W_{[6]}M = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

因此可以解出

$$A = (0,0,0,0,1,0)^{\mathrm{T}}.$$

最后, 由$A \vee B = M_d AB = (0,0,1,0,0,0)^{\mathrm{T}}$ 解出

$$B = (0,0,1,0,0,0)^{\mathrm{T}}.$$

我们得出结论, A "很不可能"是凶手, 并且B 有80% 的可能是凶手. □

从上例可以看出, 将逻辑表达式转化为矩阵形式可以使逻辑推理转化为代数方程组求解问题.

比较我们的推理和文献[79] 中的推理可以看出, 文献[79] 在模糊逻辑推理中需要建立几个非公理化的人工规则, 但是我们不需要任何额外的规则就能得到相同的结论.

下面, 我们考虑多值逻辑表达式的规范型. 与2 值逻辑一样, 我们可以证明如下定理.

定理 6.5.4 任意具有逻辑变量$P_1, \cdots, P_r$ 的k 值逻辑表达式$L(P_1, \cdots, P_r)$ 都可以唯一表示成规范型

$$L(P_1, \cdots, P_r) = M_L P_1 P_2 \cdots P_r, \tag{6.5.13}$$

其中M_L 是一个$k \times k^r$ 结构矩阵.

最后, 如前所述, 在多值逻辑中经常使用(6.5.1)~(6.5.3) 和(6.5.10). 但是蕴涵有很多种形式, 它们在某些情况下很有用, 可以强调一定的逻辑特性. 到目前为止, 我们仅使用了(6.4.10). 下面我们列出了$k = 3$ 时其他的一些蕴涵算子[105], 它们在模糊控制等中很有用.

- Zadeh: $A \rightarrow B = \neg A \vee (A \wedge B)$,

$$M_i^Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

- Lukasiewicz: $A \to B = (\neg A + B) \wedge T_0$,

$$M_i^L = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

- Mamdani: $A \to B = A \wedge B$,

$$M_i^M = M_c^3,$$

- Gaines-Rescher: $A \to B = \begin{cases} T_0, & A \leq B; \\ 0, & \text{否则}. \end{cases}$

$$M_i^{GR} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix};$$

- Reichenbach: $A \to B = \neg A + AB$,

$$M_i^R = \begin{bmatrix} 1 & 0 & 0 & 1 & 1^* & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

其中a^* 表示a 是一个四舍五入整数.

- Gödel: $A \to B = \begin{cases} T_0, & A \leq B; \\ B, & \text{否则}. \end{cases}$

$$M_i^{Gl} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix};$$

- Goguen: $A \to B = \begin{cases} T_0, & A = 0; \\ \frac{B}{A} \wedge T_0, & \text{否则}. \end{cases}$

$$M_i^{Gn} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix};$$

- Yager: $A \to B = B^A$,

$$M_i^Y = \begin{bmatrix} 1 & 0 & 0 & 1 & 1^* & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix};$$

- Kleene-Dienes: $A \to B = \neg A \vee B$,

$$M_i^{KD} = M_i^3, \quad 如(6.5.8) 所示;$$

这是本章我们所使用的.

- Wang: $A \to B = \begin{cases} T_0, & A \le B; \\ \neg A \vee B, & 否则. \end{cases}$

$$M_i^W = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

在这些蕴涵中只有M_i^Z, M_i^M 和M_i^{KD} 包含经典逻辑. M_i^{KD} 是我们在本章中使用最多的, $M_i^M = M_c$. 至于M_i^Z, 我们可以证明

$$M_i^Z = M_d M_n (I_k \otimes M_c) M_r. \tag{6.5.14}$$

对于$k > 3$, 容易算出这些蕴涵相应的结构矩阵. 于是所有其他的二元算子都可以通过它们的结构矩阵来定义.

下面, 我们将所有的二元算子推广到k 值逻辑. 这可以通过相应的结构矩阵来完成. 当然, 我们需要(6.5.3) 中定义的"否定"算子. $\sigma_8 = \wedge$ 和$\sigma_{14} = \vee$ 分别在(6.5.2) 和(6.5.1) 中定义. 对于任意k, 很容易算出它们的结构矩阵. $\sigma_0 = F_0$, $\sigma_{15} = T_0$, $\sigma_3 = \neg_1$, $\sigma_5 = \neg_2$, $\sigma_{10} = Q$ 和$\sigma_{12} = P$ 的结构矩阵也很容易构造出来.

前面已经说过蕴涵σ_{11} 有很多种定义. 对于等值, 可以用(6.5.10)算出M_e. 最后, 对于EOR (σ_6), NAND (σ_7), NOR (σ_1), NIMP (σ_4), NIIMP (σ_2), ICOD (σ_{13}), 它们的结构矩阵可以分别由(6.2.2)~(6.2.7) 算出.

现在所有的二元逻辑算子都有定义. 因此, 很自然有这样一个问题: 第4 节中的公式有多少仍然成立? 当然, 它依赖于蕴涵的定义. 在一些蕴涵下没有重言式(或$A \to A \not\Leftrightarrow T_0$) . 下面我们考虑一些含有重言式的多值逻辑的基本性质.

例 6.5.5 考虑命题6.3.4、6.3.8 和6.3.11, 对于Gödel (Gl), Lukasiewicz (L), Wang (W) 和Gaines-Reescher (GR). 利用结构矩阵, 容易检验表达式是真或假. 如表6.5.2~表6.5.4 所示.

表 6.5.2 对于某些多值逻辑检验命题6.3.4

	1)	2)	3)	4)	5)	6)	7)	8)	9)	10)
Gl	T	T	T	T	T	F	F	F	T	F
L	T	T	T	T	T	F	F	T	F	F
W	T	T	T	T	T	F	F	T	F	F
GR	T	T	T	T	T	F	F	T	F	F

这里1), 2), 3), ⋯ 表示命题6.3.4 中的第1, 第2, 第3, ⋯ 式, 即(6.3.2), (6.3.3), (6.3.4) ⋯

表 6.5.3 对于某些多值逻辑检验命题6.3.8

	1)	2)	3)	4)	5)	6)	7)	8)
Gl	T	T	F	F	T	T	T	F
L	T	T	F	F	T	T	T	F
W	T	T	F	F	T	T	T	F
GR	T	T	F	F	T	T	T	F

这里1), 2), 3), ⋯ 表示命题6.3.8 中的第1, 第2, 第3, ⋯ 式, 即(6.3.13), (6.3.14), (6.3.15) ⋯

表 6.5.4 对于某些多值逻辑检验命题6.3.11

	1)	2)	3)	4)	5)	6)	7)	8)	9)	10)	11)	12)	13)
Gl	T	T	T	T	F	T	F	T	F	T	F	T	T
L	T	T	T	T	T	T	T	T	F	F	F	F	F
W	T	T	T	T	T	T	T	T	F	F	F	F	F
GR	T	T	T	T	F	F	F	F	F	T	T	T	T

这里1), 2), 3), ⋯ 表示命题6.3.11 中的第1, 第2, 第3, ⋯ 式, 即(6.3.21), (6.3.22), (6.3.23) ⋯ □

6.6 混合值逻辑

在模糊控制中, 经常出现变量可以在不同的有限论域取值的情况. 形式上, 在一个问题中逻辑变量P_i, $i = 1, \cdots, s$ 可以从D_{k_i} 中取值, 其中k_i 可以不同. 我们给出准确的定义:

定义 6.6.1　混合值逻辑是一个四元集:

$$\mathcal{L}_m := \{D_{k_1} \cup \cdots \cup D_{k_s},\ \wedge,\ \vee,\ \neg\}.$$

在这种逻辑里逻辑变量P 从$D_{k_1} \cup \cdots \cup D_{k_s}$ 中取值. 算子$\wedge$, $\vee$, $\neg$ 分别由(6.5.2), (6.5.1) 和(6.5.3) 所定义.

容易看出, 在这种逻辑里逻辑变量可以从D_ℓ 中取值, 其中ℓ 可以和所有的k_i, $i = 1, \cdots, s$ 不同. 假设$\{0,1\}$ 是D_k, $\forall k$ 中的公共元素, 那么我们有

$$\min\{k_i \,|\, 1 \le i \le s\} \le \ell \le \sum_{i=1}^{s} k_i - 2s + 2. \tag{6.6.1}$$

注意到$D_{k_1} \cup \cdots \cup D_{k_s} \subset D_f$, 于是显然有下面的命题.

命题 6.6.2　$\mathcal{L}_m$ 是摩根代数$\{D_f,\ \wedge,\ \vee,\ \neg\}$ 的子代数.

由命题6.6.2, 模糊逻辑中的所有结果都可用于混合值逻辑$\mathcal{L}_m$, 而不需要更多的条件.

混合值逻辑的一个重要优势是可以用矩阵方法, 因此它在理论分析和数值实现方面都很方便. 本节的剩下部分主要讨论$\mathcal{L}_m$ 中算子的矩阵表示.

首先, 我们考虑二元算子的结构矩阵.

定义 6.6.3　设$P \in D_k$, $Q \in D_\mu$. 一个二元算子σ 是一个映射: $D_k \times D_\mu \to D_k \cup D_\mu$, 由它的结构矩阵$M_\sigma$ 唯一定义.

我们给出一个例子来说明.

例 6.6.4　设$P \in D_3$, $Q \in D_4$, 则

$$D_3 \cup D_4 = \{0, 1/2, 1\} \cup \{0, 1/3, 2/3, 1\} = \{0, 1/3, 1/2, 2/3, 1\} \simeq D_5.$$

注意在模糊逻辑中, D_k 中的精确值是没有意义的, 只有这些值的相对次序很重要. 因此我们可以仅根据次序将这些元素转换成向量形式. 于是我们可以构造出析取"∨" 的结构矩阵为

$$M_d^{(3,4)} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{6.6.2}$$

同理, 合取"∧" 的结构矩阵为

$$M_c^{(3,4)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}. \tag{6.6.3}$$

利用同样的技巧, 可以定义所有的二元算子的结构矩阵. 例如, 蕴涵"→" (在KD意义下) 的结构矩阵是

$$M_i^{(3,4)} = M_d^{(3,4)} M_n^3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.6.4}$$

□

称二元逻辑算子$\sigma : D_k \times D_\mu \to D_\ell$ 是对称的, 如果

$$P\sigma Q = Q\sigma P, \quad 或 \quad \sigma(P,Q) = \sigma(Q,P).$$

为了检验一个算子的对称性, 我们有

命题 6.6.5 二元算子$\sigma : D_k \times D_\mu \to D_\ell$ 是对称的, 当且仅当它的结构矩阵满足

$$M_\sigma^{(k,\mu)} W_{[\mu,k]} = M_\sigma^{(\mu,k)}. \tag{6.6.5}$$

证明 设$P \in D_k$, $Q \in D_\mu$. 对称意味着

$$M_\sigma^{(k,\mu)} PQ = M_\sigma^{(\mu,k)} QP.$$

由于

$$LHS = M_\sigma^{(k,\mu)} W_{[\mu,k]} W_{[k,\mu]} PQ = M_\sigma^{(k,\mu)} W_{[\mu,k]} QP,$$

故结论成立. □

对于混合值逻辑, 仍然有逻辑表达式的规范型定理.

定理 6.6.6 设$P_i \in D_{k_i}, i=1,\cdots,r$ 是一组混合值逻辑变量, $\cup_{i=1}^r D_{k_i} \simeq D_\ell$. 任意逻辑表达式$L(P_1,\cdots,P_r)$ 都可以唯一表示成规范型

$$L(P_1,\cdots,P_r)=M_L P_1 P_2 \cdots P_r, \tag{6.6.6}$$

其中M_L 是一个唯一的$\ell \times \prod_{i=1}^r k_i$ 矩阵, 称为L 的结构矩阵.

6.7 基于逻辑的模糊控制

回顾模糊控制系统[86, 4], 可以看出实际使用的是混合值逻辑, 包括多值逻辑作为它的特殊情形.

一个模糊(控制) 系统如图6.7.1 所示, 其中$r(t)$ 是参考输入, $u(t) \in U$ 是控制, $y(t) \in Y$ 是输出, FC 是模糊控制器, F, LBR 和DF 分别是模糊器, 基于逻辑的规则和解模糊器.

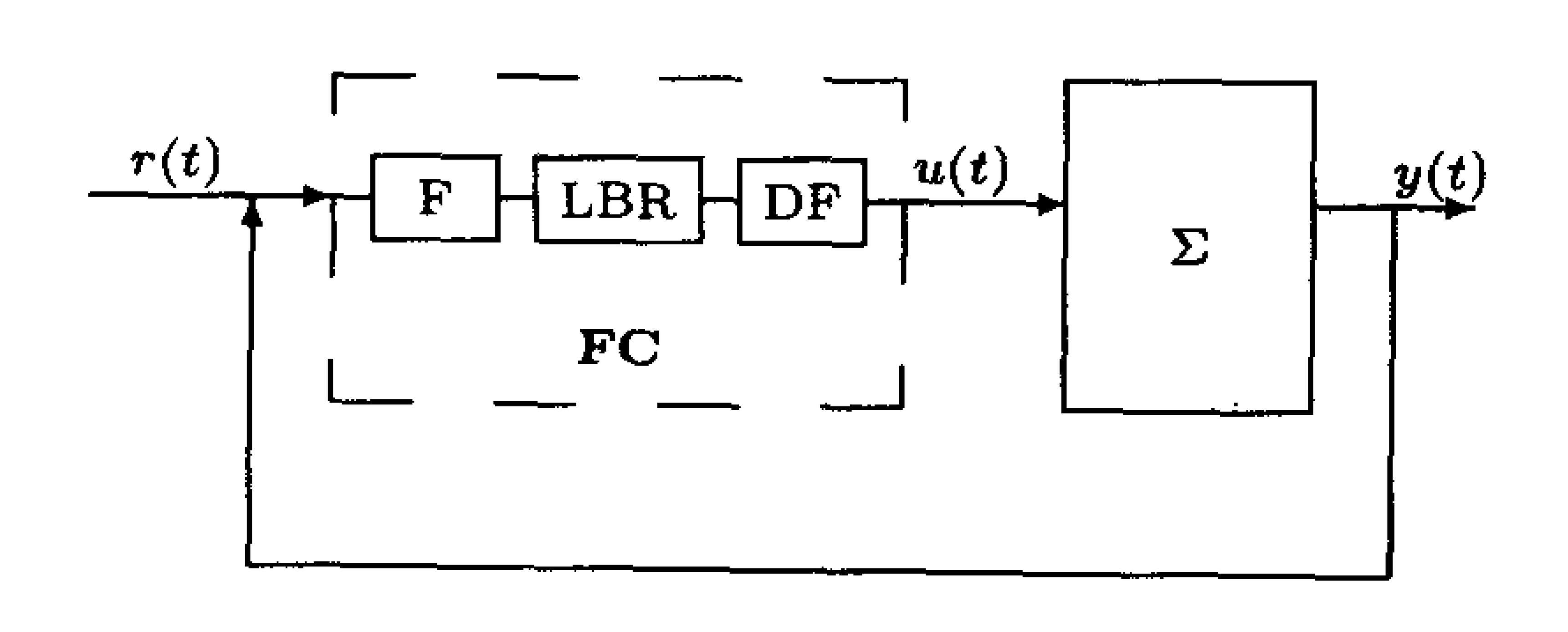

图 6.7.1　模糊控制系统

模糊控制的目的是通过基于模糊逻辑的规则得到控制$u(y)$. 我们给出一个规范的描述如下:

- 输出空间: Y (例如$Y \subset \mathbb{R}^p$);
- 输入(控制) 空间: U (例如$U \subset \mathbb{R}^m$);
- 语言命题(分别关于Y 和U):

$$\mathcal{A}_Y=\{A_1,A_2,\cdots,A_s\}; \quad \mathcal{B}_U=\{B_1,\cdots,B_t\};$$

- 相应的隶属函数:

$$\mu_{A_i}(y) \in D_f, \quad y \in Y; \quad \mu_{B_j}(u) \in D_f, \quad u \in U;$$

- 语言规则:

如果$A_1^1,\cdots,A_{s_1}^1$ 满足R_1, 那么B_1;

$$\vdots$$

如果$A_1^t,\cdots,A_{s_t}^t$ 满足R_t, 那么B_t,　　其中$A_j^i \in \mathcal{A}$;

- 与 R_i 相应的逻辑表达式是 L_i, $i=1,\cdots,t$.

语言规则可以表示成逻辑表达式如下

$$\begin{cases} L_1(A_1^1,\cdots,A_{s_1}^1) \Rightarrow B_1, \\ \vdots \\ L_t(A_1^t,\cdots,A_{s_t}^t) \Rightarrow B_t. \end{cases} \tag{6.7.1}$$

利用Mandani 蕴涵(在模糊控制中经常使用它), 第i 个规则的隶属度是

$$\mu_{R_i} = \mu_{A_1^i}(y) \wedge \cdots \wedge \mu_{A_{s_i}^i}(y) \wedge \mu_{B_i}(u), \quad i=1,\cdots,t. \tag{6.7.2}$$

于是解模糊方法用于将决定转化为行动. 例如, 使用重心法(COG) 有

$$u^{\text{crisp}} = \frac{\sum_{i=1}^{t} B_i \int \mu(i)}{\sum_{i=1}^{t} \int \mu(i)}. \tag{6.7.3}$$

或者使用中心平均法(CA) 有

$$u^{\text{crisp}} = \frac{\sum_{i=1}^{t} B_i \mu_{R_i}}{\sum_{i=1}^{t} \mu_{R_i}}. \tag{6.7.4}$$

在大多数模糊控制问题中(6.7.1) 是被查表法取代的. 实际上, 这样的规则并没有揭示输入输出的逻辑关系. 因此, 它本质上是基于经验的模糊控制.

基于逻辑的模糊控制(在后面给出) 是通过混合值模糊逻辑实现(6.7.1). 表达式都定义在$D_s \times D_t$ 上, 并且直接由y 生成反馈控制u.

我们用一个简单的例子来说明基于逻辑的模糊控制与基于经验的模糊控制相比, 说明基于逻辑的模糊控制更合理也更简洁.

例 6.7.1 [86] 考虑小车上的倒立摆. 设$e(t)$, $\dot{e}(t)$ 和$u(t)$ 分别是误差, 误差变化和推力. 输入隶属函数如图6.7.2 所示(注意到e 和$\dot{e}$ 具有不同并集的相同隶属函数), 并且输出隶属函数如图6.7.3 所示.

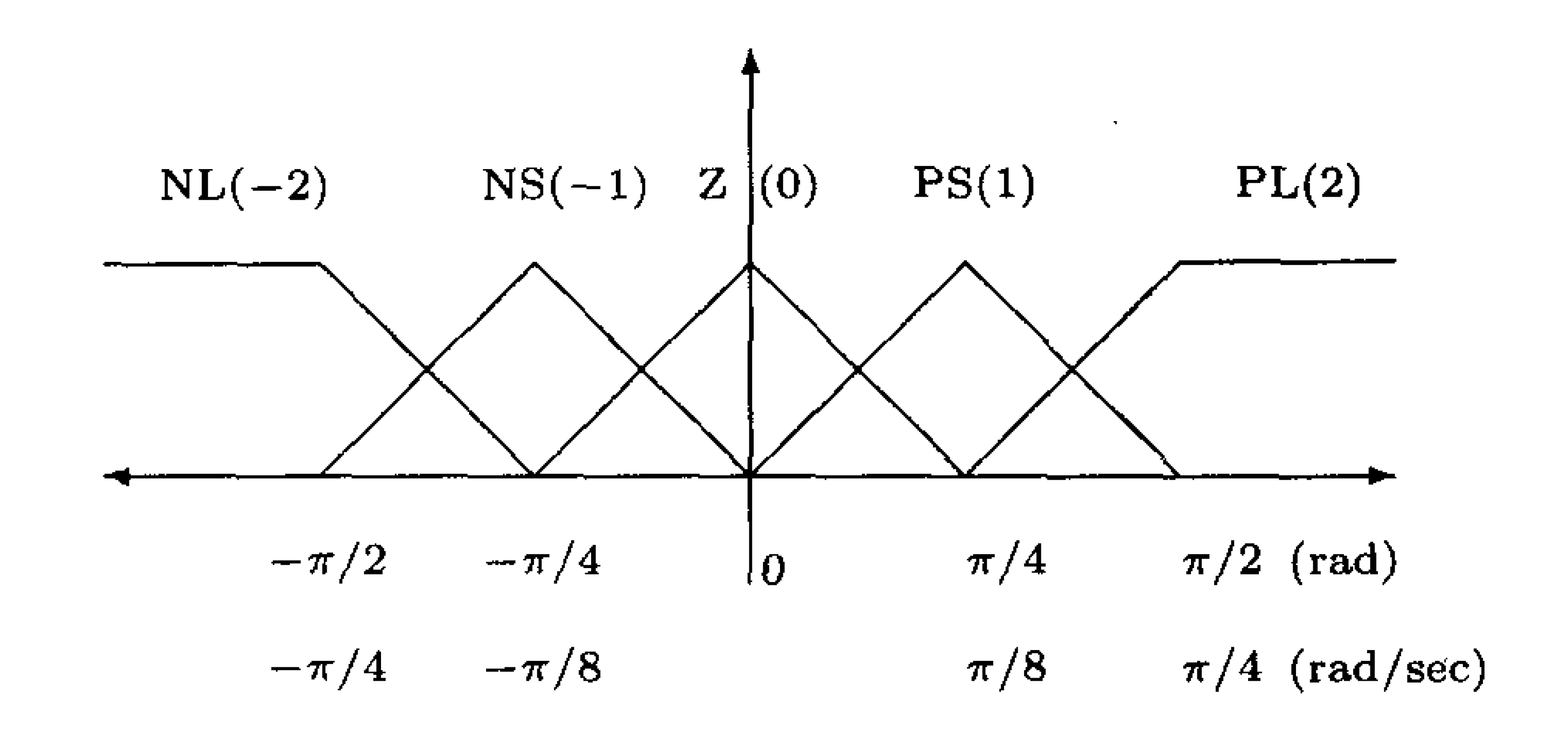

图 6.7.2　输入隶属函数

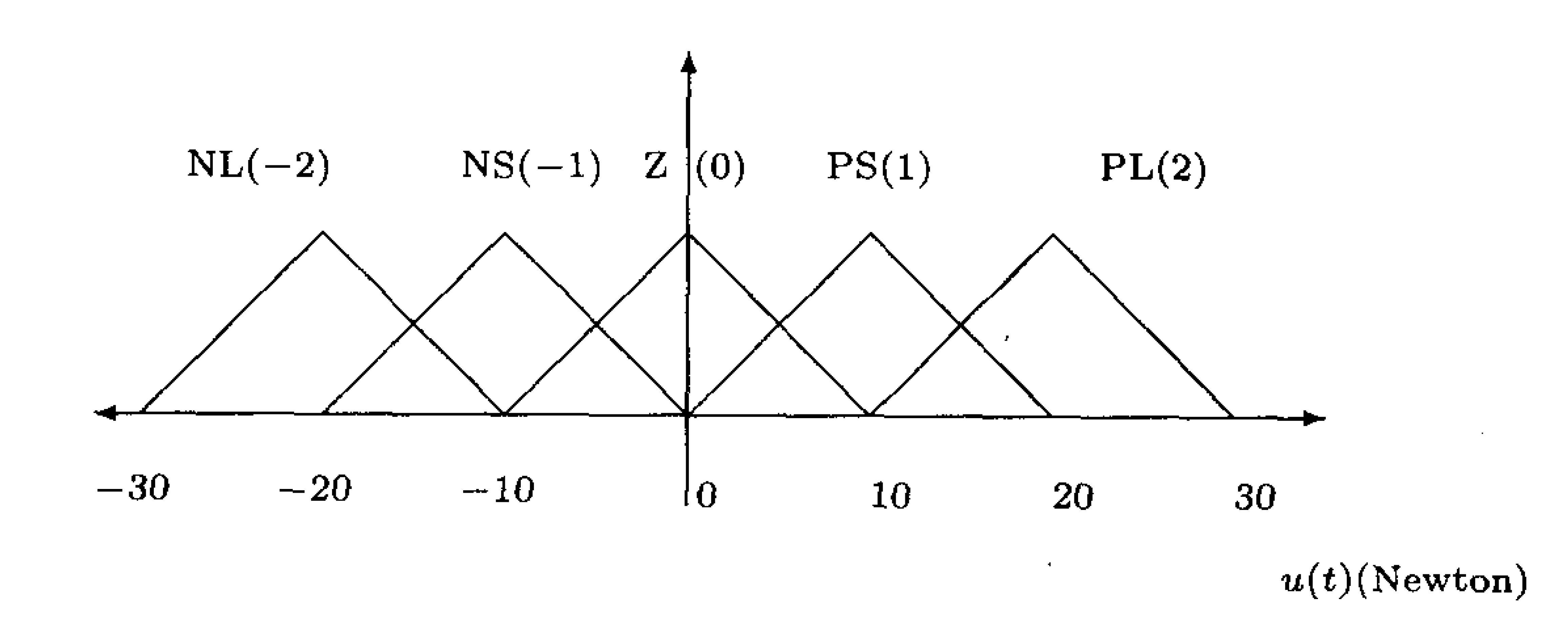

图 6.7.3　输出隶属函数

在文献[86] 中使用的规则表如表6.7.1 所示.

表 6.7.1　倒立摆的规则表

$e\backslash u\backslash \dot{e}$	−2	−1	0	1	2
−2	2	2	2	1	0
−1	2	2	1	0	−1
0	2	1	0	−1	−2
1	1	0	−1	−2	−2
2	0	−1	−2	−2	−2

注意到从表6.7.1 可以看出, $u(t)$ 与$e(t)+\dot{e}(t)$ 逻辑相关. 形式上, 我们正则化(改变比例) e 和$\dot{e}$ 为

$$E = 4e/\pi; \quad \dot{E} = 8\dot{e}/\pi.$$

对于$E+\dot{E}$ 和$u(t)$, 我们分别用A_i 和B_i, $i=1,2,3,4,5$ 表示NL, NS, Z, PS, PL. 可以得到(6.7.1) 的逻辑规则是

$$\mathcal{B} = \neg\mathcal{A}. \tag{6.7.5}$$

下面让我们看看如何设计基于逻辑的模糊控制.

对$E+\dot{E}$ 使用经典隶属函数, 见图6.7.4.

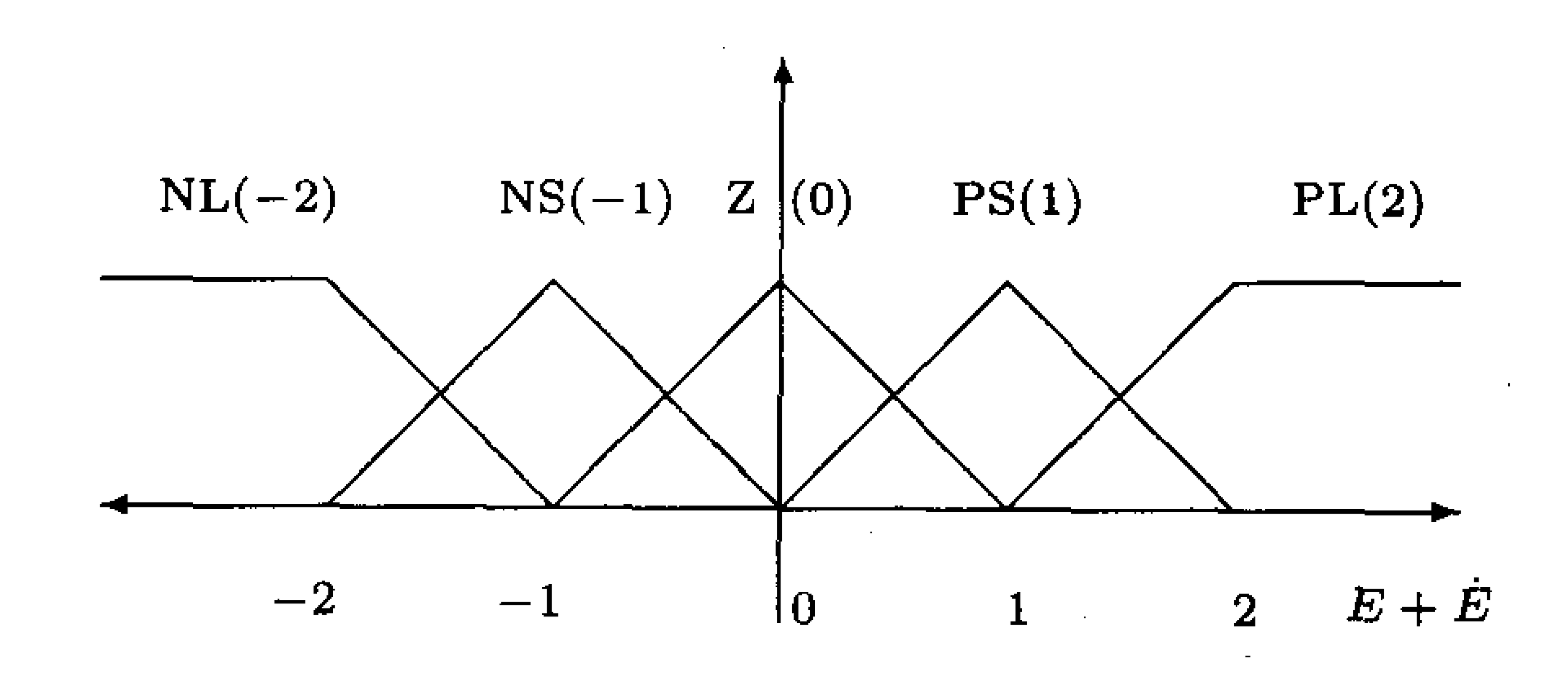

图 6.7.4 $E+\dot{E}$ 的隶属函数

下面我们选择e 和$\dot{e}$ 的一些特殊值来比较关于基于经验(E-B) 和关于基于逻辑(L-B) 方法的控制$u(t)$ (见表6.7.2). □

表 6.7.2 比较通过基于经验方法和通过基于逻辑方法的模糊控制

e		$-\frac{3\pi}{8}$			0			$\frac{5\pi}{16}$		
$\dot{e}$		$-\frac{\pi}{32}$	$\frac{3\pi}{32}$	$\frac{9\pi}{32}$	$-\frac{\pi}{32}$	$\frac{3\pi}{32}$	$\frac{9\pi}{32}$	$-\frac{\pi}{32}$	$\frac{3\pi}{32}$	$\frac{9\pi}{32}$
COG	E-B	16.84	3.18	−10	8.68	−6.82	−18.06	−5	−20	−20
	L-B	16.67	2.5	−10	8.33	−7.5	−18.33	−5	−20	−20
CA	E-B	16.82	6.82	−6.82	3.18	−6.82	−20	−10	−20	−20
	L-B	17.5	7.5	−7.5	2.5	−7.5	−20	−10	−20	−20

注 (1) 从例6.7.1 可以看出基于逻辑的模糊控制只需要更少的计算.

(2) 这种方法具有广泛的应用. 例如, 文献[86] 中的大部分例子都可以使用这种方法.

(3) 显然, 基于逻辑的方法可以用于更一般的逻辑关系, 其中包括基于经验的方法无法解决的情况.

(4) 这里提出的基于逻辑的模糊控制应当说只是一种初步的尝试, 这方面目前还没有什么成熟的结果或方法. 需要进一步的深入研究和探讨.

6.8 注释与参考

数理逻辑是很多大学的一门基础课程. 本章关于数理逻辑的一些基本概念可参考一些经典教材, 例如文献[58].

模糊逻辑是经典逻辑的推广, 它与经典逻辑的最大区别是, 在经典逻辑中逻辑变量只能取$\{0,1\}$两个离散值, 而在模糊逻辑中逻辑变量可以取$[0,1]$中的任意值. 参考例如文献[79].

多值逻辑介于经典逻辑和模糊逻辑之间. 在多值逻辑中逻辑变量可以取$[0,1]$间预先指定的有限个值. 参见文献[52].

关于基于逻辑的智能系统中的问题, 一本很好的参考书是文献[96].

模糊控制是对无法精确描述的系统依多个模糊规则进行控制. 它比较符合现实世界. 同时在实践中得到许多应用. 但它在理论上还比较初等, 并且, 模糊逻辑未得到充分应用. 关于模糊控制可参考文献[123], [102], [104].

本章内容部分基于文献[31], [32].

第 7 章　几何和代数中的半张量积方法

本章我们讨论半张量积在微分几何和近世代数方面的一些应用. 在数学问题中, 经常会出现多指标, 在这些情况下, 半张量积可以用来简化记号和计算. 我们还将讨论几何和代数中的一些典型例子. 首先, 我们给出微分几何中一些重要概念的半张量积计算公式, 包括联络及其相关运算、曲率算子及黎曼曲率张量等的计算. 然后讨论有穷维代数的结构矩阵, 对于李代数, 给出结构矩阵应满足的半张量积形式的代数方程. 从而给出一些低维李代数的完全刻画. 最后, 对相对论中应用广泛的张量场缩并公式, 利用半张量积给出简单证明.

7.1　联络及其运算

联络是几何和物理中的一个重要概念. 本节利用半张量积给出一些联络的基本计算公式的简单矩阵表示. 首先给出联络的定义:

定义 7.1.1　*设 $f, g \in V(M)$ 是流形 M 上的两个光滑(C^∞) 向量场. 一个$\mathbb{R}$-双线性映射∇: $V(M) \times V(M) \to V(M)$ 如果满足*

(1)

$$\nabla_{rf} sg = rs\nabla_f g, \quad r, s \in \mathbb{R}; \tag{7.1.1}$$

(2)

$$\nabla_{hf} g = h\nabla_f g, \quad \nabla_f (hg) = L_f(h)g + h\nabla_f g, \quad h \in C^\infty(M). \tag{7.1.2}$$

则称该映射为联络.

根据上述定义, 只要在基向量上联络定义好了, 则对所有向量场的联络也就定义好了. 因此, 首先要定义局部坐标下联络对基向量的运算. 在局部坐标卡x下, 基向量上的运算由下式决定:

$$\nabla_{\frac{\partial}{\partial x_i}} \left(\frac{\partial}{\partial x_j} \right) = \sum_{k=1}^{n} \gamma_{ij}^k \frac{\partial}{\partial x_k},$$

其中γ_{ij}^k 称为Christoffel 记号. 我们称矩阵

$$\Gamma = \begin{bmatrix} \gamma_{11}^1 & \cdots & \gamma_{1n}^1 & \cdots & \gamma_{n1}^1 & \cdots & \gamma_{nn}^1 \\ \vdots & & \vdots & & \vdots & & \vdots \\ \gamma_{11}^n & \cdots & \gamma_{1n}^n & \cdots & \gamma_{n1}^n & \cdots & \gamma_{nn}^n \end{bmatrix}$$

为Christoffel 矩阵.

下面我们给出两个向量场之间的联络的矩阵表示:

命题 7.1.2 设 $f = \sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i}$, $g = \sum_{j=1}^{n} g_j \frac{\partial}{\partial x_j}$, 记其向量形式为$f = (f_1, f_2, \cdots, f_n)^{\mathrm{T}}$, $g = (g_1, g_2, \cdots, g_n)^{\mathrm{T}}$, 则

$$\nabla_f g = Dgf + \Gamma fg. \tag{7.1.3}$$

证明 根据(7.1.1)~(7.1.2) 联络的定义, 我们可以算出

$$\begin{aligned}\nabla_f g &= \sum_{i=1}^{n} f_i \left[\sum_{j=1}^{n} L_{\frac{\partial}{\partial x_i}} g_j \frac{\partial}{\partial x_j} + \sum_{j=1}^{n}\sum_{k=1}^{n} g_j \gamma_{ij}^k \frac{\partial}{\partial x_k} \right] \\ &= Dg \ltimes f + \Gamma \ltimes f \ltimes g.\end{aligned} \tag{7.1.4}$$

因此我们就得到(7.1.3). □

设$y = y(x)$ 是另一个局部坐标卡, 现在我们在新坐标卡下推导矩阵Γ 的公式. 记$\tilde{\Gamma}$ 和$\tilde{\gamma}_{ij}^k$ 为新坐标卡下相应的Γ 和γ_{ij}^k, 于是我们有

引理 7.1.3 在新坐标卡y 下, 我们有如下公式:

$$\begin{aligned}\begin{bmatrix} \tilde{\gamma}_{ij}^1 \\ \vdots \\ \tilde{\gamma}_{ij}^n \end{bmatrix} &= \begin{bmatrix} \frac{\partial^2 x_1}{\partial y_j \partial y_1} & \cdots & \frac{\partial^2 x_1}{\partial y_j \partial y_n} \\ \vdots & & \vdots \\ \frac{\partial^2 x_n}{\partial y_j \partial y_1} & \cdots & \frac{\partial^2 x_n}{\partial y_j \partial y_n} \end{bmatrix} \left[\frac{\partial x_1}{\partial y_i}, \quad \cdots, \quad \frac{\partial x_n}{\partial y_i} \right]^{\mathrm{T}} \\ &\quad + \Gamma \ltimes \left[\frac{\partial x_1}{\partial y_i}, \quad \cdots, \quad \frac{\partial x_n}{\partial y_i} \right]^{\mathrm{T}} \ltimes \left[\frac{\partial x_1}{\partial y_j}, \quad \cdots \quad, \frac{\partial x_n}{\partial y_j} \right]^{\mathrm{T}}\end{aligned} \tag{7.1.5}$$

证明 令

$$f = \frac{\partial}{\partial y_i} = \sum_{s=1}^{n} \frac{\partial}{\partial x_s} \frac{\partial x_s}{\partial y_i},$$

$$g = \frac{\partial}{\partial y_j} = \sum_{t=1}^{n} \frac{\partial}{\partial x_t} \frac{\partial x_t}{\partial y_j}.$$

回顾γ 的定义, 我们有

$$\sum_{k=1}^{n} \tilde{\gamma}_{ij}^k \frac{\partial}{\partial y_k} = \nabla_f g.$$

应用(7.1.3) 到上式, 于是(7.1.5) 成立. □

定理 7.1.4 在新坐标卡y 下, 有

$$\tilde{\Gamma} = D^2xDx + \Gamma \ltimes Dx(I \otimes Dx). \tag{7.1.6}$$

证明 直接计算有

$$D^2x \ltimes Dx = \begin{bmatrix} \sum_{s=1}^n \frac{\partial^2 x_1}{\partial y_s \partial y_1}\frac{\partial x_s}{\partial y_1} & \cdots & \sum_{s=1}^n \frac{\partial^2 x_1}{\partial y_s \partial y_n}\frac{\partial x_s}{\partial y_1} \\ \vdots & \ddots & \vdots \\ \sum_{s=1}^n \frac{\partial^2 x_n}{\partial y_s \partial y_1}\frac{\partial x_s}{\partial y_1} & \cdots & \sum_{s=1}^n \frac{\partial^2 x_n}{\partial y_s \partial y_n}\frac{\partial x_s}{\partial y_1} \end{bmatrix.$$
$$\begin{matrix} \cdots & \sum_{s=1}^n \frac{\partial^2 x_1}{\partial y_s \partial y_1}\frac{\partial x_s}{\partial y_n} & \cdots & \sum_{s=1}^n \frac{\partial^2 x_1}{\partial y_s \partial y_n}\frac{\partial x_s}{\partial y_n} \\ \ddots & \vdots & \ddots & \vdots \\ \cdots & \sum_{s=1}^n \frac{\partial^2 x_n}{\partial y_s \partial y_1}\frac{\partial x_s}{\partial y_n} & \cdots & \sum_{s=1}^n \frac{\partial^2 x_n}{\partial y_s \partial y_n}\frac{\partial x_s}{\partial y_n} \end{bmatrix}.$$

如果我们将它的列由(ij) 按照索引$(11,\cdots,1n,\ \cdots\ ,n1,\cdots,nn)$ 排列, 那么它的第(ij) 列就是(7.1.5) 右边的第一项.

记J_i 为Dx 的第i 列, 则

$$\Gamma \ltimes Dx = (\Gamma \ltimes J_1, \Gamma \ltimes J_2, \cdots, \Gamma \ltimes J_n).$$

我们也有$I \otimes Dx = \mathrm{diag}(J,\cdots,J)$, 这里$J=(J_1,\cdots,J_n)$, 因此

$$\begin{aligned} &\Gamma \ltimes Dx \ltimes (I \otimes Dx) \\ &= (\Gamma \ltimes J_1 \ltimes J_1, \cdots, \Gamma \ltimes J_1 \ltimes J_n, \cdots, \Gamma \ltimes J_n \ltimes J_1, \cdots, \Gamma \ltimes J_n \ltimes J_n). \end{aligned}$$

容易看出上式中的第(ij) 列是(7.1.5) 右边的第二项. □

注 如果我们用右半张量积, (7.1.6) 也可以写成

$$\tilde{\Gamma} = D^2x \ltimes Dx + (\Gamma \ltimes Dx) \rtimes Dx. \tag{7.1.7}$$

由于左半张量积和右半张量积之间不满足结合律, 需要注意运算的次序. 为了避免可能的混淆, 我们尽量不用右半张量积.

设M 是一个黎曼流形, 它的黎曼度量由对称矩阵$G=(g_{ij})_{n\times n}$ 决定. 黎曼几何的基本定理说, M 上存在唯一确定的黎曼联络[1]. 而且, 这个联络的Christoffel 记号由G 依下式计算:

$$\gamma_{ij}^k = \frac{1}{2}\sum_{s=1}^n g^{ks}\left(\frac{\partial g_{si}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_s} + \frac{\partial g_{js}}{\partial x_i}\right), \tag{7.1.8}$$

其中g^{ij} 表示G^{-1} 的第(i,j) 元素.

众所周知[10], 在唯一确定了联络的黎曼流形中有

$$[f,g]=\nabla_f g-\nabla_g f. \tag{7.1.9}$$

称Christoffel 矩阵是对称的, 如果

$$\gamma_{ij}^k=\gamma_{ji}^k, \quad \forall i,j,k. \tag{7.1.10}$$

于是我们有

定理 7.1.5 如果流形N 具有对称Christoffel 矩阵的联络, 则(7.1.9) 成立.

证明 容易看出, 如果Christoffel 矩阵Γ 对称, 则$\Gamma W_{[n]}=\Gamma$, 也就意味着

$$\Gamma fg=\Gamma gf, \quad \forall f,g\in V(N).$$

利用(7.1.3) 有

$$\nabla_f g-\nabla_g f=Dgf-Dfg=[f,g].$$

□

由(7.1.8) 可以看出对于黎曼流形, Christoffel 矩阵对称. 因此(7.1.9) 成立, 这说明定理7.1.5 给出了一个更加一般的结果.

在利用计算机进行计算或进行公式推导时, 矩阵形式常常是有效且方便的. 直接验证可以看出(7.1.8) 的矩阵形式是

$$\Gamma=\frac{1}{2}G^{-1}\left(DG+DGW_{[n]}-(DV_r(G))^{\mathrm{T}}\right). \tag{7.1.11}$$

一个相关话题就是测地线. 我们可以将测地线方程[1] 表示成矩阵形式. 曲线$r(t)$ 是M 上的测地线, 当且仅当

$$\ddot{r}_t=\Gamma\ltimes\dot{r}_t^2. \tag{7.1.12}$$

如果$r(t)$ 在局部坐标下表示成$r(t)=(x_1(t),\cdots,x_n(t))^{\mathrm{T}}$, 那么(7.1.12) 按照分量形式可以写成

$$\begin{bmatrix}\ddot{x}_1\\ \vdots\\ \ddot{x}_n\end{bmatrix}=\Gamma\begin{bmatrix}\dot{x}_1\\ \vdots\\ \dot{x}_n\end{bmatrix}^2.$$

最后, 我们考虑曲率算子和黎曼曲率张量.

定义 7.1.6 [10] (1) 曲率算子$R(X,Y)$ 是由向量场X 和Y 决定的算子, 并且对每个向量场Z 指定一个新的C^∞ 向量场$R(X,Y)\cdot Z$ 为

$$R(X,Y)\cdot Z=\nabla_X(\nabla_Y Z)-\nabla_Y(\nabla_X Z)-\nabla_{[X,Y]}Z. \tag{7.1.13}$$

(2) 黎曼曲率张量是一个4 阶的C^∞ 协变张量, 定义为

$$\mathcal{R}(X,Y,Z,W)=(R(X,Y)\cdot Z,W). \tag{7.1.14}$$

现在让我们计算曲率算子的结构矩阵.

给定局部坐标卡$x=(x_1,\cdots,x_n)$ 和$E_i:=\frac{\partial}{\partial x_i}$. 分别记三个向量场$X=(\alpha_1,\cdots,\alpha_n)$, $Y=(\beta_1,\cdots,\beta_n)$, $Z=(\gamma_1,\cdots,\gamma_n)$, 则[10]

$$R(X,Y)\cdot Z=\sum_{i=1}^n\sum_{j=1}^n\sum_{k=1}^n\alpha_i\beta_j\gamma_k R(E_i,E_j)\cdot E_k. \tag{7.1.15}$$

由于

$$\begin{aligned}\nabla_{E_i}(\nabla_{E_j}E_k) &= \nabla_{E_i}\left(\sum_{t=1}^n\gamma_{jk}^t E_t\right)\\ &= \sum_{t=1}^n L_{E_i}(\gamma_{jk}^t)E_t+\sum_{t=1}^n\gamma_{jk}^t\sum_{\ell=1}^n\gamma_{it}^\ell E_\ell\\ &= \sum_{t=1}^n\left(L_{E_i}(\gamma_{jk}^t)+\sum_{\ell=1}^n\gamma_{i\ell}^t\gamma_{jk}^\ell\right)E_t.\end{aligned} \tag{7.1.16}$$

同理

$$\nabla_{E_j}(\nabla_{E_i}E_k)=\sum_{t=1}^n\left(L_{E_j}(\gamma_{ik}^t)+\sum_{\ell=1}^n\gamma_{j\ell}^t\gamma_{ik}^\ell\right)E_t. \tag{7.1.17}$$

将它们应用到(7.1.13) 和(7.1.15), 我们可以按如下方式构造结构矩阵: 定义

$$m_{ijk}^t:=\left(L_{E_i}(\gamma_{jk}^t)+\sum_{\ell=1}^n\gamma_{i\ell}^t\gamma_{jk}^\ell\right)-\left(L_{E_j}(\gamma_{ik}^t)+\sum_{\ell=1}^n\gamma_{j\ell}^t\gamma_{ik}^\ell\right),$$

则曲率算子R 的结构矩阵是

$$M_R=\left(m_{ijk}^t\right), \tag{7.1.18}$$

其中元素按照$\binom{t}{ijk}$ 矩阵形式(参见定义1.2.1) 排列. 因此

$$R(X,Y)\cdot Z=M_R XYZ. \tag{7.1.19}$$

注意前面的推导是形式的, 要使(7.1.19) 成立, 必须R 对X, Y, Z 关于函数也是3 线性的, 即对任何光滑函数f 有$R(fX,Y,Z)=fR(X,Y,Z)$ 等, 参见文献[10].

下面考虑黎曼曲率张量. 首先注意到, 对于一个双线性型$p^{\mathrm{T}}Gq$, 我们有

$$p^{\mathrm{T}}Gq=\sum_{i,j}g_{ij}p_iq_j=V_r^{\mathrm{T}}(G)pq=V_c^{\mathrm{T}}(G)qp,$$

对于黎曼曲率张量, 注意到

$$\mathcal{R}(X,Y,Z,W)=W^{\mathrm{T}}GR(X,Y)\cdot Z.$$

并利用上述双线性表达式, 有

$$\mathcal{R}(X,Y,Z,W)=V_c^{\mathrm{T}}(G)M_{\mathcal{R}}XYZW. \tag{7.1.20}$$

7.2　有限维代数的结构分析

本节我们利用半张量积分析有限维代数的结构. 首先, 我们考虑一个有限维代数的矩阵表示.

定义 7.2.1 [67]　一个n 维代数是$\mathbb{R}$ 上的n 维向量空间$\mathcal{L}$, 并且它的乘法$*$ 满足分配律, 即

$$\begin{aligned}&(aX+bY)*Z=a(X*Z)+b(Y*Z),\\&Z*(aX+bY)=a(Z*X)+b(Z*Y),\quad X,Y,Z\in\mathcal{L},\ a,b\in\mathbb{R}.\end{aligned} \tag{7.2.1}$$

定义 7.2.2　设$\{e_1,\cdots,e_n\}$ 是$\mathcal{L}$ 的一组基底, 并假设

$$e_i*e_j=\sum_{k=1}^{n}\alpha_{ij}^ke_k,\quad i,j=1,\cdots,n. \tag{7.2.2}$$

则称α_{ij}^k 为结构常数.

称矩阵

$$M_{\mathcal{L}}=\begin{bmatrix}\alpha_{11}^1&\cdots&\alpha_{1n}^1&\cdots&\alpha_{n1}^1&\cdots&\alpha_{nn}^1\\\alpha_{11}^2&\cdots&\alpha_{1n}^2&\cdots&\alpha_{n1}^2&\cdots&\alpha_{nn}^2\\\vdots&&\vdots&&\vdots&&\vdots\\\alpha_{11}^n&\cdots&\alpha_{1n}^n&\cdots&\alpha_{n1}^n&\cdots&\alpha_{nn}^n\end{bmatrix} \tag{7.2.3}$$

为代数$\mathcal{L}$ (在乘法$*$ 下) 的结构矩阵.

固定这组基底, 并记$X = \sum_{i=1}^{n} x_i e_i, Y = \sum_{i=1}^{n} y_i e_i$ 等, 将它们简化为系数向量$X = (x_1, \cdots, x_n)^{\mathrm{T}}$, $Y = (y_1, \cdots, y_n)^{\mathrm{T}}$ 等, 那么显然有

命题 7.2.3 设$Z = X * Y$, 则得到系数向量形式的Z 为

$$Z = M_{\mathcal{L}} \ltimes X \ltimes Y = M_{\mathcal{L}} XY. \tag{7.2.4}$$

我们前面提过$\ltimes$ 是普通矩阵乘法的推广, 我们可以省略乘法符号$\ltimes$. 在后面将省略所有的$\ltimes$. 因此当两个矩阵不满足等维数条件时, 则一定是左半张量积.

一个代数的所有性质由它的结构矩阵决定. 下面我们研究一些基本的性质.

定义 7.2.4 给定一个代数$\mathcal{L}$.

(1) $\mathcal{L}$ 称为对称的, 如果

$$X * Y = Y * X, \quad \forall X, Y \in \mathcal{L}; \tag{7.2.5}$$

(2) $\mathcal{L}$ 称为反对称的, 如果

$$X * Y = -Y * X, \quad \forall X, Y \in \mathcal{L}; \tag{7.2.6}$$

(3) $\mathcal{L}$ 满足结合律, 如果

$$(X * Y) * Z = X * (Y * Z), \quad \forall X, Y, Z \in \mathcal{L}. \tag{7.2.7}$$

命题 7.2.5 给定一个n 维代数$\mathcal{L}$.

(1) $\mathcal{L}$ 对称, 当且仅当

$$M_{\mathcal{L}}(W_{[n]} - I_{n^2}) = 0; \tag{7.2.8}$$

(2) $\mathcal{L}$ 反对称, 当且仅当

$$M_{\mathcal{L}}(W_{[n]} + I_{n^2}) = 0; \tag{7.2.9}$$

(3) $\mathcal{L}$ 满足结合律, 当且仅当

$$M_{\mathcal{L}}(M_{\mathcal{L}} \otimes I_n - I_n \otimes M_{\mathcal{L}}) = 0. \tag{7.2.10}$$

为了证明这个命题, 我们需要一个引理, 它本身也很有用.

回忆命题2.4.3, 设V_1, V_2 是两个维数分别为n_1, n_2 的向量空间. 张量积空间$V = V_1 \otimes V_2$ 是由$\{x \otimes y \,|\, x \in V_1, y \in V_2\}$ 生成的空间. 进一步, 设$\Phi : V_1 \times V_2 \to W$ 是一个双线性映射, 那么存在唯一的$\Psi : V \to W$ 使得图7.2.1 交换.

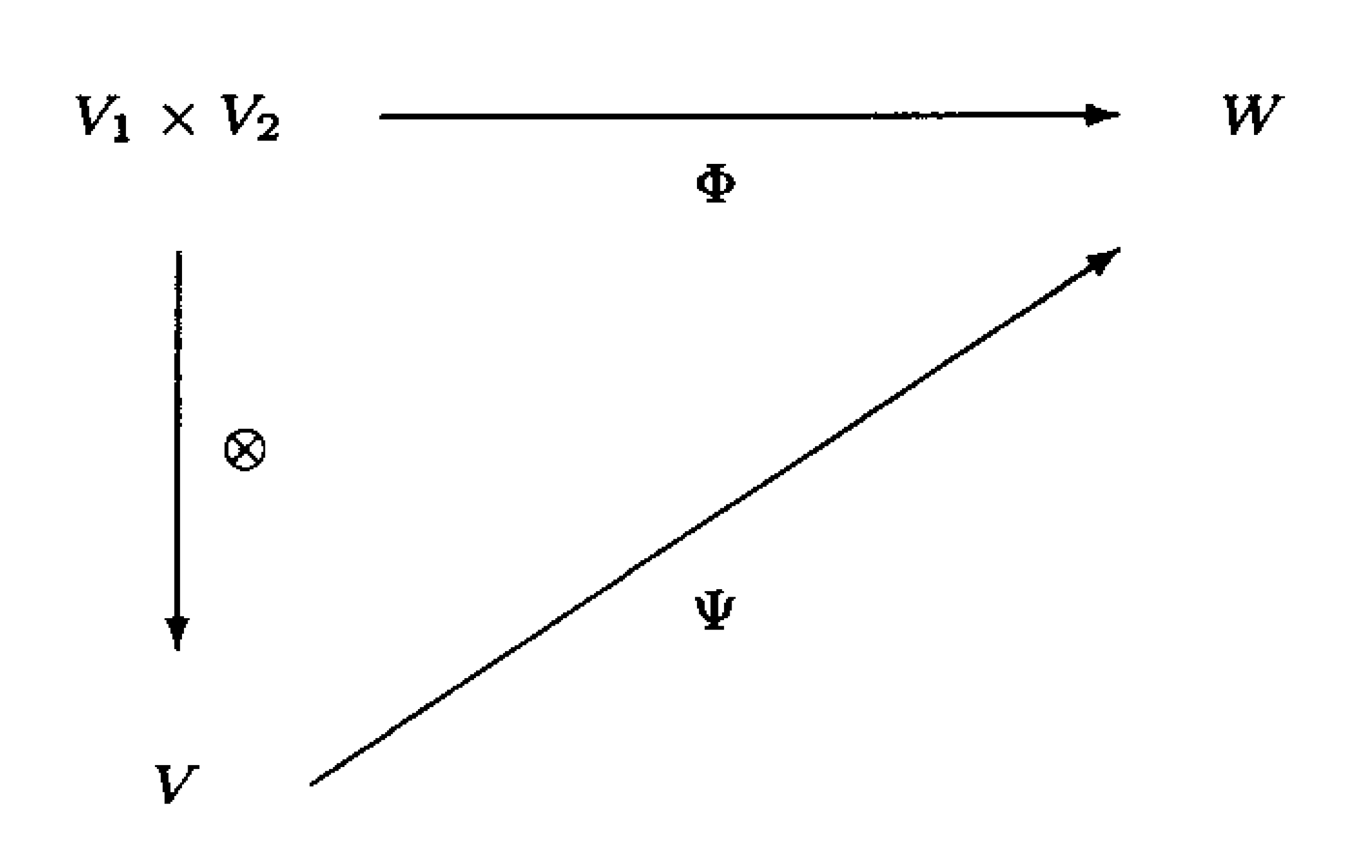

图 7.2.1　张量积空间

同理, 可以定义具有k 重线性映射Φ 的k 个向量空间的张量积.

设V_i, $i=1,\cdots,k$ 是维数和基底分别为n_i 和$\{d^i_1,\cdots,d^i_{n_i}\}$ 的k 个向量空间, 且W 是具有固定基底的m 维向量空间. 在这组固定的基底下, 多线性映射Φ : $V_1\times\cdots\times V_k\to W$ 具有结构矩阵M_Φ, 它是一个$m\times n$ 矩阵, 这里$n=\prod_{i=1}^k n_i$. 张量积空间$V=\Pi_{i=1}^k V_i$ 具有基底

$$\begin{aligned}\{e_1,\cdots,e_n\}:=\{e^1_{i_1}\otimes e^2_{i_1}\otimes\cdots\otimes e^k_{i_k}\mid i_k=1,\cdots,n_k,\\ i_{k-1}=1,\cdots,n_{k-1},\cdots,i_1=1,\cdots,n_1\}.\end{aligned}\tag{7.2.11}$$

于是存在唯一诱导映射$\Psi:V\to W$ 使得Φ 和$\Psi\circ\otimes$ 可交换. 利用这些记号, 我们有下面的引理.

引理 7.2.6　(1) 在(7.2.11) 中定义的基底下, Ψ 与Φ 的结构矩阵相同, 即

$$M_\Psi=M_\phi.\tag{7.2.12}$$

(2) 如果

$$\Phi(X_1,\cdots,X_k)=0,\quad \forall X_i\in V_i,$$

那么

$$\Psi(Y)=0,\quad \forall Y\in V.$$

证明　(1) 从映射的构造可以看出

$$\Psi(e^1_{i_1}\otimes\cdots\otimes e^k_{i_k})=\Phi(e^1_{i_1},\cdots,e^k_{i_k}),\quad \forall 1\le i_t\le n_t,\ t=1,\cdots,k.$$

注意

$$\{e^1_{i_1}\otimes\cdots\otimes e^k_{i_k}\mid 1\le i_t\le n_t, t=1,\cdots,k\}$$

构成了V 的一组基底, 即(7.2.12) 在V 的基底下成立. 由于$\Psi:V\to W$ 是线性映射, 于是(7.2.12) 对所有的V 都成立.

(2) 它是(1) 的直接结果, 因为它是(1) 中$M_\phi=0$ 的情形. □

命题7.2.5 的证明 我们只证明(7.2.10). (7.2.8) 和(7.2.9) 的证明类似. 利用(7.2.4), (7.2.7) 可以写成矩阵形式

$$MM(XY)Z = MX(MYZ), \quad \forall X, Y, Z \in \mathcal{L}.$$

由结合律我们有

$$M^2XYZ = M(XM)YZ = M(I_n \otimes M)XYZ, \quad \forall X, Y, Z \in \mathcal{L}.$$

于是

$$(M^2 - M(I_n \otimes M))XYZ = 0, \quad \forall X, Y, Z \in \mathcal{L}. \tag{7.2.13}$$

注意到尽管集合

$$S = \{XYZ \mid X, Y, Z \in \mathcal{L}\}$$

不是向量空间, 由引理7.2.6, 结论仍成立. □

下面, 我们考虑代数是李代数时的情形.

命题 7.2.7 一个代数$\mathcal{L}$ 是李代数, 当且仅当它的结构矩阵满足: (i) 反对称性, 即(7.2.9); (ii) Jacobi 恒等式, 即

$$M^2(I_{n^2} + W_{[n,n^2]} + W_{[n^2,n]}) = 0. \tag{7.2.14}$$

证明 一个代数是李代数, 当且仅当它反对称, 并且满足Jacobi 恒等式:

$$(X * Y) * Z + (Y * Z) * X + (Z * X) * Y = 0, \quad \forall X, Y, Z \in \mathcal{L}. \tag{7.2.15}$$

(i) 来自命题7.2.5.

利用结构矩阵, (7.2.15) 可以表示成

$$M^2(XYZ + YZX + ZXY) = 0. \tag{7.2.16}$$

利用换位矩阵的性质, 我们有

$$W_{[n,n^2]}XYZ = YZX, \quad W_{[n^2,n]}XYZ = ZXY.$$

将它们代入到(7.2.16), 并利用类似命题7.2.5 的证明, (7.2.15) 成立. □

例 7.2.8 $\mathbb{R}^3$ 中的叉乘定义为: 设$X = x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$, $Y = y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$, 则

$$X \times Y = \det \begin{bmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix}.$$

于是得到它的结构矩阵为

$$M=\begin{bmatrix}0&0&0&0&0&1&0&-1&0\\0&0&-1&0&0&0&1&0&0\\0&1&0&-1&0&0&0&0&0\end{bmatrix}. \tag{7.2.17}$$

可以验证

$$M\left(I_9+W_{[3]}\right)=0,$$

且

$$M^2\left(I_{27}+W_{[3.9]}+W_{[9,3]}\right)=0.$$

因此$\mathbb{R}^3$ 中的叉乘是一个李代数. □

为构造$gl(n,\mathbb{R})$ 的结构矩阵, 我们先给出两个矩阵乘积的列展开的公式, 它本身也很有用.

引理 7.2.9　设$A\in M_{m\times n}$, $B\in M_{n\times p}$, 则

$$V_c(AB)=\Psi_{mnp}V_c(A)V_c(B), \tag{7.2.18}$$

其中

$$\Psi_{mnp}=\begin{bmatrix}I_m\otimes(\delta_1^p\delta_1^n)^{\mathrm{T}}&I_m\otimes(\delta_1^p\delta_2^n)^{\mathrm{T}}&\cdots&I_m\otimes(\delta_1^p\delta_n^n)^{\mathrm{T}}\\I_m\otimes(\delta_2^p\delta_1^n)^{\mathrm{T}}&I_m\otimes(\delta_2^p\delta_2^n)^{\mathrm{T}}&\cdots&I_m\otimes(\delta_2^p\delta_n^n)^{\mathrm{T}}\\\vdots&\vdots&&\vdots\\I_m\otimes(\delta_p^p\delta_1^n)^{\mathrm{T}}&I_m\otimes(\delta_p^p\delta_2^n)^{\mathrm{T}}&\cdots&I_m\otimes(\delta_p^p\otimes\delta_n^n)^{\mathrm{T}}\end{bmatrix}, \tag{7.2.19}$$

这里δ_i^p (δ_i^n) 是I_p (I_n) 的第i 列.

证明　回顾表3.2.1, 有

$$V_c(AB)=(I_p\otimes A)V_c(B).$$

因此我们需要算出$I_p\otimes A$. 直接计算有

$$I_p\otimes A=\Psi_{mnp}V_c(A).$$

下面我们考虑$gl(n,\mathbb{R})$ 的结构矩阵, 并以一个例子来说明. 为了记号的方便, 我们记$\Psi_n:=\Psi_{nnn}$.

例 7.2.10 考虑$gl(n,\mathbb{R})$. 选择$\{M_{I,J},\ I=1,\cdots,n;\ J=1,\cdots,n\}$ 作为一组基底, 其中$M_{I,J}$ 的元素定义为

$$(M_{I,J})_{i,j}=\begin{cases}1, & i=I \text{ 且 } j=J,\\ 0, & \text{其他}.\end{cases}$$

现在我们构造李括号

$$[A,B]=AB-BA$$

的结构矩阵, 于是我们有

$$V_c(AB)=\Psi_n V_c(A)V_c(B),$$

且

$$V_c(BA)=\Psi_n V_c(B)V_c(A)=\Psi_n W_{[n^2]}V_c(A)V_c(B),$$

因此

$$V_c([A,B])=(\Psi_n-\Psi_n W_{[n^2]})V_c(A)V_c(B),$$

这表示$gl(n,\mathbb{R})$ 的结构矩阵是

$$M=\Psi_n(I_{n^4}-W_{[n^2]}). \tag{7.2.20}$$

我们可以检验M 满足(7.2.9) 和(7.2.14). □

下面, 我们想问这样一个问题: 什么是($\mathbb{R}$ 上) n 维代数集合的拓扑结构?

注意到一个代数由它的结构矩阵唯一决定, 我们可以依其结构矩阵的维数将$\mathbb{R}^{n\times n^2}$ 空间的普通欧氏空间拓扑赋予n 维代数. 由于结构矩阵是依赖于坐标的, 而由坐标变换引起的形式不同的结构矩阵是等价的. 因此在这个等价意义下的商拓扑是一个合适的描述. 设$M_{\mathcal{L}_n}$ 是关于基底$\{e_1\}$ 的结构矩阵, 存在另一组基底$\{e_2\}$, 并且对于某个$T\in GL(n,\mathbb{R})$, $e_2=e_1T$. 于是对于任两个向量$V_1=e_1X_1$ 和$V_2=e_1X_2$, 在新基底下可以表示成$\tilde{X}_i=T^{-1}X_i$, $i=1,2$, 我们有

$$\begin{aligned}&T^{-1}M_{\mathcal{L}_n}X_1X_2=\tilde{M}_{\mathcal{L}_n}T\tilde{X}_1T\tilde{X}_2\\ &=\tilde{M}_{\mathcal{L}_n}T^{-1}(I_n\otimes T^{-1})X_1X_2=\tilde{M}_{\mathcal{L}_n}(T^{-1}\otimes T^{-1})X_1X_2.\end{aligned}$$

于是有

$$\tilde{M}_{\mathcal{L}_n}=T^{-1}M_{\mathcal{L}_n}(T\otimes T). \tag{7.2.21}$$

将这个结果用于两个李代数, 我们有

命题 7.2.11 两个n 维李代数$\mathcal{L}_n^1$ 和$\mathcal{L}_n^2$ 等价, 当且仅当存在$T \in GL(n,\mathbb{R})$ 使得

$$M_{\mathcal{L}_n^1} = T^{-1} M_{\mathcal{L}_n^2}(T \otimes T). \tag{7.2.22}$$

因此我们可以给出n 维代数集合的商拓扑, 记为

$$\mathbb{R}^{n \times n^2}/GL(n,\mathbb{R}).$$

现在我们考虑李代数的集合. 首先, 我们考虑$n = 2$ 时的情形. 为了保证反对称性, 一个2 维李代数的结构矩阵是

$$M_{\mathcal{L}_2} = \begin{bmatrix} 0 & a & -a & 0 \\ 0 & b & -b & 0 \end{bmatrix}. \tag{7.2.23}$$

直接计算有

$$M_{\mathcal{L}_2}^2 = \begin{bmatrix} 0 & 0 & -ab & a^2 & ab & -a^2 & 0 & 0 \\ 0 & 0 & -b^2 & ab & b^2 & -ab & 0 & 0 \end{bmatrix}. \tag{7.2.24}$$

$$I_8 + W_{[2,4]} + W_{[4,2]} = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}. \tag{7.2.25}$$

于是

$$M_{\mathcal{L}_2}^2 (I_8 + W_{[2,4]} + W_{[4,2]}) = 0.$$

这样我们有下面的命题:

命题 7.2.12 任意2 维反对称代数是李代数.

接着, 我们考虑3 维的情形. 为了保证反对称, 它的结构矩阵必须是

$$M_{\mathcal{L}_3} = \begin{bmatrix} 0 & a & d & -a & 0 & g & -d & -g & 0 \\ 0 & b & e & -b & 0 & h & -e & -h & 0 \\ 0 & c & f & -c & 0 & i & -f & -i & 0 \end{bmatrix}. \tag{7.2.26}$$

通过计算机我们可以算出

$$M_{\mathcal{L}_3}^2(I_{27}+W_{[3,9]}+W_{[9,3]}),$$

它是一个3×27 矩阵. 幸运地是, 它只有很少几个不同的非零元素, 它们是

$$m_{1,6}=m_{1,16}=m_{1,22}=-m_{1,8}=-m_{1,12}=-m_{1,20}=bg+gf-ah-di;$$

$$m_{2,6}=m_{2,16}=m_{2,22}=-m_{2,8}=-m_{2,12}=-m_{2,20}=ae-bd+hf-ei;$$

$$m_{3,6}=m_{3,16}=m_{3,22}=-m_{3,8}=-m_{3,12}=-m_{3,20}=af+bi-cd-ch.$$

于是我们得出结论

定理 7.2.13 一个3 维代数是李代数, 当且仅当它的结构矩阵如(7.2.26) 所示, 并且其元素满足下列方程组:

$$\begin{cases}bg+gf-ah-di=0,\\ ae-bd+hf-ei=0,\\ af+bi-cd-ch=0.\end{cases}\tag{7.2.27}$$

例 7.2.14 根据定理7.2.13, 我们可以构造很多3 维李代数.

(1) 设

$$a=b=d=f=h=i=0,\quad c=g=1,\quad e=-1.$$

容易检验这是(7.2.27) 的一个解. 实际上, 这个结构生成了含有标准叉乘的代数$\mathbb{R}^3$.

(2) 为了看看如何轻松地得到(7.2.27) 的一些非零特解, 我们可以将它转化成矩阵形式.

$$\begin{bmatrix}-h & g & 0\\ e & -d & 0\\ f & i & -d-h\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}=\begin{bmatrix}di-gf\\ ei-hf\\ 0\end{bmatrix}.\tag{7.2.28}$$

找到(7.2.27) 非零解的一个(不是必要的) 方法是选择d,e,f,g,h,i, 使得(7.2.28) 的系数矩阵是非奇异的. 然后解出相应的解a,b,c 就可以了. 例如, 如果我们选择$d=-e=1$, $f=-g=2$, $h=-i=3$, 那么我们有

$$\begin{bmatrix}-3 & -2 & 0\\ -1 & -1 & 0\\ 2 & -3 & -4\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}=\begin{bmatrix}1\\ -3\\ 0\end{bmatrix}.$$

它的解是$a=-7$, $b=10$, $c=-11$, 即我们有下面的李代数

$$\mathcal{L}_3=\{\alpha I+\beta J+\gamma K\,|\,\alpha,\beta,\gamma\in\mathbb{R}\},$$

它的乘法$*$ 满足

$$I*I=J*J=K*K=0,$$

$$I*J=-J*I=-7I+10J-11K,$$

$$I*K=-K*I=I-J+2K,$$

$$J*K=-K*J=-2I+3J-3K.$$

□

下面, 我们考虑 4 维的情形. 为了保证反对称, 它的4×16 的结构矩阵$M_{\mathcal{L}_4}=(m_{ij})$ 要满足

$$
\begin{cases}
m_{i,j}=0, \quad j=1,6,11,16;\\
m_{i,2}=-m_{i,5}:=x_i,\\
m_{i,3}=-m_{i,9}:=x_{4+i},\\
m_{i,4}=-m_{i,13}:=x_{8+i},\\
m_{i,7}=-m_{i,10}:=x_{12+i},\\
m_{i,8}=-m_{i,14}:=x_{16+i},\\
m_{i,12}=-m_{i,15}:=x_{20+i},\\
\qquad\qquad i=1,2,3,4.
\end{cases}
\tag{7.2.29}
$$

使用Matlab, 一个合适的程序可以将(7.2.14) 转化为

$$
\begin{cases}
-x_1x_{14}+x_2x_{13}-x_4x_{21}-x_5x_{15}+x_7x_{13}+x_8x_{17}-x_9x_{16}=0,\\
x_1x_6-x_2x_5-x_4x_{22}-x_6x_{15}+x_7x_{14}+x_8x_{18}-x_{10}x_{16}=0,\\
x_1x_7+x_2x_{15}-x_3x_5-x_3x_{14}-x_4x_{23}+x_8x_{19}-x_{11}x_{16}=0,\\
x_1x_8+x_2x_{16}-x_4x_5-x_4x_{14}-x_4x_{24}+x_7x_{16}-x_8x_{15}\\
\quad +x_8x_{20}-x_{12}x_{16}=0,\\
-x_1x_{18}+x_2x_{17}+x_3x_{21}-x_5x_{19}-x_9x_{20}+x_{11}x_{13}+x_{12}x_{17}=0,\\
x_1x_{10}-x_2x_9+x_3x_{22}-x_6x_{19}-x_{10}x_{20}+x_{11}x_{14}+x_{12}x_{18}=0,\\
x_1x_{11}+x_2x_{19}-x_3x_9-x_3x_{18}+x_3x_{23}-x_7x_{19}+x_{11}x_{15}\\
\quad -x_{11}x_{20}+x_{12}x_{19}=0,\\
x_1x_{12}+x_2x_{20}+x_3x_{24}-x_4x_9-x_4x_{18}-x_8x_{19}+x_{11}x_{16}=0,\\
-x_1x_{22}-x_5x_{23}+x_6x_{17}+x_7x_{21}-x_9x_{24}-x_{10}x_{13}+x_{12}x_{21}=0,\\
-x_2x_{22}+x_5x_{10}-x_6x_9+x_6x_{18}-x_6x_{23}+x_7x_{22}-x_{10}x_{14}\\
\quad -x_{10}x_{24}+x_{12}x_{22}=0,\\
-x_3x_{22}+x_5x_{11}+x_6x_{19}-x_7x_9-x_{10}x_{15}-x_{11}x_{24}+x_{12}x_{23}=0,\\
-x_4x_{22}+x_5x_{12}+x_6x_{20}+x_7x_{24}-x_8x_9-x_8x_{23}-x_{10}x_{16}=0,\\
x_1x_{21}-x_5x_{17}+x_9x_{13}-x_{13}x_{18}-x_{13}x_{23}+x_{14}x_{17}\\
\quad +x_{15}x_{21}-x_{17}x_{24}+x_{20}x_{21}=0,\\
x_2x_{21}-x_6x_{17}+x_{10}x_{13}-x_{14}x_{23}+x_{15}x_{22}-x_{18}x_{24}+x_{20}x_{22}=0,\\
x_3x_{21}-x_7x_{17}+x_{11}x_{13}+x_{14}x_{19}-x_{15}x_{18}-x_{19}x_{24}+x_{20}x_{23}=0,\\
x_4x_{21}-x_8x_{17}+x_{12}x_{13}+x_{14}x_{20}+x_{15}x_{24}-x_{16}x_{18}\\
\quad -x_{16}x_{23}=0.
\end{cases}
\tag{7.2.30}
$$

定理 7.2.15 一个4 维代数是李代数，当且仅当它的结构矩阵满足(7.2.29)，并且$x_1, \cdots, x_{24}$ 满足方程组(7.2.30).

例 7.2.16 (1) 考虑$gl(2,\mathbb{R})$. 根据例7.2.10, 我们选择$\{e_1, e_2, e_3, e_4\}$ 作为基底，这里

$$e_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad e_3 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad e_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

于是由简单计算即可得到结构矩阵为

$$M = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(2) 实际上, (7.2.30) 有很多非零解. 例如，令

$$x_i = 0, \quad i > 8,$$

于是(7.2.30) 变成

$$\begin{cases} x_1x_6 - x_2x_5 = 0, \\ x_1x_7 - x_3x_5 = 0, \\ x_1x_8 - x_4x_5 = 0. \end{cases}$$

因此它是解，当且仅当

$$x_1 : x_2 : x_3 : x_4 = x_5 : x_6 : x_7 : x_8.$$

例如, $x_1 = 1$, $x_2 = -1$, $x_3 = 2$, $x_4 = -2$, $x_5 = -1$, $x_6 = 1$, $x_7 = -2$, $x_8 = 2$. 因此我们有一个李代数

$$\mathcal{L} = \{aI + bJ + cK + dH \mid a, b, c, d \in \mathbb{R}\},$$

并且它的乘法$*$ 满足

$$\begin{cases} I * I = J * J = K * K + H * H = 0, \\ I * J = -J * I = -I + J - 2K + 2H, \\ I * K = -K * J = I - J + 2K - 2H, \\ I * H = -H * I = J * K = -K * J = J * H \\ \quad = -H * J = K * H = -H * K = 0. \end{cases}$$

□

最后, 我们考虑n 维的情形. 假设一个n 维代数的结构矩阵可以表示成

$$M_{\mathcal{L}_n} = \begin{bmatrix} W_{11} & \cdots & W_{1n} & \cdots & W_{n1} & \cdots & W_{nn} \end{bmatrix},$$

其中每个W_{ij} 是一个列, 且w_{ij}^s 是它的元素. 于是

$$M_{\mathcal{L}_n^2} = \left[\sum_{s=1}^{n} w_{ij}^s W_{sk} \,\middle|\, i,j,k = 1,\cdots,n \right]. \tag{7.2.31}$$

注意(7.2.31) 中的列是由指标(i,j,k) 按照字母序排列的, 即k 先从1 到n, 然后j 从1 到n, 然后i 从1 到n. 当将它乘以$W_{[n^2,n]}$ ($W_{[n,n^2]}$) 时, 它就从按照指标(i,j,k) 排序变成按照指标(j,k,i) $((k,i,j))$ 排序. 考虑到这点, 我们有

$$\sum_{s=1}^{n} w_{ij}^s W_{sk} + \sum_{s=1}^{n} w_{jk}^s W_{si} + \sum_{s=1}^{n} w_{ki}^s W_{sj} = 0.$$

再考虑到反对称, 我们只要考虑i, j, k 互不相同的情形, 于是我们有

$$\begin{aligned} &\sum_{s<k} w_{ij}^s W_{sk} - \sum_{s>k} w_{ij}^s W_{ks} + \sum_{s<i} w_{jk}^s W_{si} - \sum_{s>i} w_{jk}^s W_{is} + \sum_{s<j} w_{ki}^s W_{sj} \\ &\quad - \sum_{s>j} w_{ki}^s W_{js} = 0, \quad 1 \le i < j < k \le n. \end{aligned} \tag{7.2.32}$$

这样我们得到

$$\frac{k!n}{3!(k-3)!}$$

个独立的方程. 于是李代数的集合可以描述成多项式方程集合的簇[59].

最后, 我们考虑代数的可逆性.

定义 7.2.17　给定一个代数$\mathcal{L}$, 它的乘法是$*$. $\mathcal{L}$ 是可逆的, 如果

(i) 存在一个单位元, 使得

$$e * X = X, \quad 且 \quad X * e = X, \quad \forall X \in \mathcal{L}; \tag{7.2.33}$$

(ii) 对任意的$X \neq 0$, 存在唯一的$X^{-1} \in \mathcal{L}$ 使得

$$X * X^{-1} = e. \tag{7.2.34}$$

此外, 如果$\mathcal{L}$ 对称(即$*$ 可交换), 称之为数字域.

注意到由于$0 \neq e \in \mathcal{L}$, 它可以作为基底的一个元素. 实际上, 由e 张成的1 维子空间是$\mathbb{R}$. 为了方便, 当存在单位元时, 我们总是按将e 作为基底的第一个分量方式选择$\mathcal{L}$ 的基底. 即自然基底$B = \{e, e_2, e_3, \cdots, e_n\}$, 其中$e = (1, 0, \cdots, 0)^{\mathrm{T}}$.

现在假设基底B 固定, 于是我们有相应的结构矩阵M. 让我们看看如何验证定义7.2.17 中的条件(i) 和(ii). 分解结构矩阵M 为

$$M = \begin{bmatrix} M_1 & M_2 & \cdots & M_n \end{bmatrix},$$

其中M_i 是$n \times n$ 矩阵. 下面的结果是定义的一个直接结论.

引理 7.2.18 存在单位元$e = (1, 0, \cdots, 0)^{\mathrm{T}}$, 当且仅当$M_1 = I_n$, $M_j^1 = \delta_j^n$, $j = 1, \cdots, n$, 其中M_j^1 是M_j 的第1列, δ_j^n 是I_n 的第j 列.

证明

$$e * X = MeX = [M_1, \cdots, M_n]\,\delta_1^n X = M_1 X.$$

因为我们要求

$$M_1 X = X, \quad \forall X \in \mathbb{R}^n.$$

故$M_1 = I_n$.

类似地

$$X * e = MW_{[n]}eX = \left[M_1^1, \cdots, M_n^1\right] X = X, \quad \forall X \in \mathbb{R}^n,$$

这说明$\left[M_1^1, \cdots, M_n^1\right] = I_n$. □

下面, 我们考虑(ii). 我们需要做的是: 对于任意$X \neq 0$, 存在唯一的Y 使得

$$MX * Y = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{7.2.35}$$

于是, 不难证明

引理 7.2.19 假设$\mathcal{L}$ 有单位元$e = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^{\mathrm{T}}$, 那么任意的$X \neq 0$ 都有唯一逆元X^{-1}, 当且仅当

$$\det(MX) \neq 0, \quad X \neq 0. \tag{7.2.36}$$

下面, 我们给出一些例子来说明.

例 7.2.20 考虑复数集合$\mathbb{C}$, 它的基底是$\{1, i\}$, 那么它的结构矩阵是

$$M_{\mathbb{C}} = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

我们检验对称性和(i), (ii).

(1) 对于对称性, 注意到

$$W_{[2]} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

容易验证

$$M_{\mathbb{C}} W_{[2]} = M_{\mathbb{C}}.$$

(2) 对于(i), 设$X = (\alpha, \beta)^{\mathrm{T}}$ 是任意的, 则

$$\begin{aligned} X * e &= M_{\mathbb{C}} X e = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \end{aligned}$$

(3)

$$\det (M_{\mathbb{C}} X) = \det \left(\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \right) = \alpha^2 + \beta^2.$$

因此行列式是0, 当且仅当$X = 0$. □

例 7.2.21 考虑四元数, 它的标准基底是$\{1, I, J, K\}$, 结构矩阵是

$$M_Q = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \tag{7.2.37}$$

容易验证(i). 对于(ii), 设$X = \begin{bmatrix} a & b & c & d \end{bmatrix}^{\mathrm{T}} \neq 0$, 我们有

$$M_Q X = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix},$$

且

$$\begin{aligned} E: &= \det(M_Q X) \\ &= a^4 + b^4 + c^4 + d^4 + 2(a^2b^2 + a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 + c^2d^2) \\ &= (a^2 + b^2 + c^2 + d^2)^2 > 0. \end{aligned} \tag{7.2.38}$$

因此四元数是可逆的. 最后

$$X^{-1} = (M_Q X)^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} := \frac{1}{E} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix},$$

其中

$$\alpha = \det\left(\begin{bmatrix} a & -d & c \\ d & a & -b \\ -c & b & a \end{bmatrix}\right) = a^3 + a(b^2 + c^2 + d^2);$$

$$\beta = -\det\left(\begin{bmatrix} b & -d & c \\ c & a & -b \\ d & b & a \end{bmatrix}\right) = -b^3 - b(a^2 + c^2 + d^2);$$

$$\gamma = \det\left(\begin{bmatrix} b & a & c \\ c & d & -b \\ d & -c & a \end{bmatrix}\right) = -c^3 - c(a^2 + b^2 + d^2);$$

$$\delta = -\det\left(\begin{bmatrix} b & a & -d \\ c & d & a \\ d & -c & b \end{bmatrix}\right) = -d^3 - d(a^2 + b^2 + c^2):$$

它不是域, 因为它不是可交换的. □

注 一般来说, 如果$\mathcal{L}$ 是对称的, 即它对乘法是可交换的, 则$X * X^{-1} = e$ 意味着$X^{-1} * X = e$. 但在一般情况下, 定义7.2.17 只规定了右逆. 可交换是左、右逆均存在且相等的充分条件.

下面给出一个左右逆存在且相等的一个较弱的要求.

命题 7.2.22 如果$\mathcal{L}$ 是一个有右(左) 单位元的结合代数, 它的每个非零元均有右(左)逆元, 则右(左)单位元也是左(右)单位元, 每个非零元的右(左)逆元也是其左(右)逆元.

证明 显见, 这里所有非零元对乘法构成群, 而本命题来自群的基本性质. □

由上述命题引出的一个自然问题是, 什么时候一个代数是结合代数? 我们有下述条件.

命题 7.2.23 n 维代数$\mathcal{L}$ 是一个结合代数, 当且仅当, 它的结构矩阵满足如下条件

$$M_{\mathcal{L}}^2 = M_{\mathcal{L}}(I_n \otimes M_{\mathcal{L}}). \tag{7.2.39}$$

证明

$$(X * Y) * Z = M_{\mathcal{L}}(M_{\mathcal{L}}XY)Z = M_{\mathcal{L}}^2XYZ,$$

$$X * (Y * Z) = M_{\mathcal{L}}X(M_{\mathcal{L}}YZ) = M_{\mathcal{L}}XM_{\mathcal{L}}YZ = M_{\mathcal{L}}(I_n \otimes M_{\mathcal{L}})XYZ.$$

于是结论成立. □

例 7.2.24 考虑例7.2.21 中的四元数. 容易验证, 其结构矩阵M_Q 满足(7.2.39), 因此, 例7.2.21 中得到的逆元既是左逆也是右逆. □

验证已知结果不是这个一般描述的主要目的, 问题是: 我们可以找到其他的满足这个一般描述的数字域吗? 下面进一步讨论这个问题.

首先, 我们考虑维数是2 的情形. 假设$\{1, \xi\}$ 是一组基底, 它使得

$$\mathcal{F}_2 := \{a + b\xi \,|\, a, b \in \mathbb{R}\}$$

是一个数字域. 我们需要定义乘法, 显然它要满足对称性和(i). 结构矩阵应该是

$$M = \begin{bmatrix} 1 & 0 & 0 & \alpha \\ 0 & 1 & 1 & \beta \end{bmatrix}. \tag{7.2.40}$$

考虑条件(ii). 对于任意的$X=(x,y)^{\mathrm{T}}$ 我们有

$$\det(MX)=x^2+\beta xy-\alpha y^2.$$

为了保证$\det(MX)>0, \forall X\neq 0$, 我们需要

$$\Delta=\beta^2+4\alpha<0. \tag{7.2.41}$$

于是我们有下面的定理.

定理 7.2.25 一个$\mathbb{R}$ 上的2 维代数是数字域, 当且仅当对于一个标准基底$(1,\xi)$, 它的结构矩阵形如(7.2.40), 其中$\alpha,\beta\in\mathbb{R}$ 满足

$$|\beta|<2\sqrt{-\alpha}.$$

注 由(7.2.40) 可以看出下面两个事实:

(1)

$$\xi^2=\alpha+\beta\xi. \tag{7.2.42}$$

(2) 只要$X=(x,y)^{\mathrm{T}}\neq 0$, 它的逆元是

$$(x+y\xi)^{-1}=\frac{1}{x^2+\beta xy-\alpha y^2}\left[(x+\beta y)-y\xi\right]. \tag{7.2.43}$$

例 7.2.26 定义一个2 维代数$\mathcal{J}$ 为

$$\mathcal{J}=\{a+b\mathbf{j}\,|\,\alpha,\beta\in\mathbb{R}\},$$

它的乘法的结构矩阵是

$$M=\begin{bmatrix}1&0&0&-1\\0&1&1&1\end{bmatrix}. \tag{7.2.44}$$

根据定理7.2.25, 它是数字域. 利用(7.2.42) 有

$$\mathbf{j}^2=-1+\mathbf{j}.$$

现在$\mathcal{J}$ 中两个数的乘积, 例如

$$(3+2\mathbf{j})(2-\mathbf{j})=6+\mathbf{j}-2\mathbf{j}^2=8-\mathbf{j}.$$

考虑除式$(3+2\mathbf{j})/(2-\mathbf{j})$, 利用(7.2.43) 我们有

$$\frac{1}{2-\mathbf{j}}=\frac{1}{3}(1+\mathbf{j}).$$

因此

$$\frac{3+2\mathbf{j}}{2-\mathbf{j}}=\frac{1}{3}+\frac{7}{3}\mathbf{j}.$$

□

命题 7.2.27　任意两个2 维域同构.

证明　设$\mathcal{J}=\{a+b\mathbf{j}\,|\,a,b\in\mathbb{R}\}$ 是一个2 维域, 并且它的结构矩阵是

$$M_{\mathcal{J}}=\begin{bmatrix}1&0&0&\alpha\\0&1&1&\beta\end{bmatrix}.$$

定义线性映射$\Phi:\mathcal{J}\to\mathbb{C}$ 为

$$1\mapsto 1,\quad \mathbf{j}\mapsto\frac{\beta}{2}+\frac{\mathbf{i}}{2}\sqrt{-(4\alpha+\beta^2)}.$$

容易验证Φ 是同构. 因此, 任何一个2 维域均与复数域同构.　□

例 7.2.28　给定一个3 维代数$\mathcal{L}_3$. 假设它是一个数字域, 那么对于一个标准基底$\{e,I,J\}$, 其中e 是单位元, 并且考虑对称性, 它的结构矩阵应该是

$$M_{\mathcal{L}_3}=\begin{bmatrix}1&0&0&0&a&d&0&d&g\\0&1&0&1&b&e&0&e&h\\0&0&1&0&c&f&1&f&i\end{bmatrix}.\tag{7.2.45}$$

现在只需检验它何时满足(ii). 设$X=(x,y,z)^{\mathrm{T}}\in\mathcal{V}$, 则

$$\det\left(M_{\mathcal{L}_3}X\right)=\det\left(\begin{bmatrix}x&ay+dz&dy+gz\\y&x+by+ez&ey+hz\\z&cy+fz&z+fy+iz\end{bmatrix}\right)=x^3+LDT(x),$$

其中$LDT(x)$ 表示x 中的低次项. 显然, 它不是正定的. 于是我们得出结论, 没有3 维的数字域.　□

魏尔斯特拉斯在1861 年证明了: 实际上,实系数的线性结合代数. 如果要求它对乘法交换, 则只有实数与复数两种. 我们有下面一个几乎可逆的代数.

例 7.2.29 设$\mathcal{L}$ 是一个具有单位元的4 维可交换代数, 它的结构矩阵是

$$M_{\mathcal{L}}=\begin{bmatrix}1&0&0&0&0&-1&0&0&0&0&1&0&0&0&0&-1\\0&1&0&0&1&0&0&0&0&0&0&-1&0&0&-1&0\\0&0&1&0&0&0&0&1&1&0&0&0&0&1&0&0\\0&0&0&1&0&0&-1&0&0&-1&0&0&1&0&0&0\end{bmatrix}. \tag{7.2.46}$$

设$X=(x,y,z,w)^{\mathrm{T}}\in\mathbb{R}^4$, 直接计算有

$$\det(M_{\mathcal{L}}X)=(x^2-z^2)^2+(y^2-w^2)^2+2(xy+zw)^2+2(xw+yz)^2.$$

因此, $x+iy+jz+kw$ 可逆, 当且仅当

$$\begin{bmatrix}x\\y\end{bmatrix}\neq\pm\begin{bmatrix}z\\w\end{bmatrix}.$$

$\mathcal{L}$ 除零测集外可逆.

$$\left\{(x,y,z,w)^{\mathrm{T}}\in\mathbb{R}^4\,\middle|\,(x,y)=\pm(z,-w)\right\}.$$

□

最后, 我们考虑在一般情况下如何计算det(MX) (如(7.2.36) 所示) .

记一个n 维代数的结构矩阵为

$$M_n=\begin{bmatrix}W_{11}&\cdots&W_{1n}&\cdots&W_{n1}&\cdots&W_{nn}\end{bmatrix}, \tag{7.2.47}$$

其中$W_{ij}\in\mathbb{R}^n$ 是它的列. 于是我们可以给出计算det(M_nX) 的公式.

设$k_1,\cdots,k_n$ 是非负整数, 且$k_1+k_2+\cdots+k_n=n$, 于是我们可以定义$P(k_1,\cdots,k_n)$ 为

$$\left\{\underbrace{1,\cdots,1}_{k_1},\underbrace{2,\cdots,2}_{k_2},\cdots,\underbrace{n,\cdots,n}_{k_n}\right\}$$

的置换的集合. 利用(7.2.47), 我们有

$$\det(M_nX)=\sum_{k_1+\cdots+k_n=n}\mu(k_1,\cdots,k_n)x_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}, \tag{7.2.48}$$

其中

$$\mu(k_1,\cdots,k_n)=\sum_{(\alpha_1,\cdots,\alpha_n)\in P(k_1,\cdots,k_n)}\det\left(\begin{bmatrix}W_{11}&\cdots&W_{1\alpha_1}&\cdots&W_{n\alpha_n}\end{bmatrix}\right).$$

例如, 设$n=3$, $k_1=0$, $k_2=2$, $k_3=1$, 那么

$$P(k_1,k_2,k_3)=\mathrm{Perm}\{2,2,3\}=\{(2,2,3),(2,3,2),(3,2,2)\}.$$

这里$\mathrm{Perm}\{S\}$ 表示S 中所有元素的不同置换.

于是, 单项$x_2^2x_3$ 的系数是

$$\begin{aligned}\mu(0,2,3) &= \det\left(\begin{bmatrix}W_{21} & W_{22} & W_{33}\end{bmatrix}\right)+\det\left(\begin{bmatrix}W_{21} & W_{32} & W_{23}\end{bmatrix}\right)\\ &\quad +\det\left(\begin{bmatrix}W_{31} & W_{22} & W_{2,3}\end{bmatrix}\right).\end{aligned}$$

7.3 张量场的缩并

半张量积是处理矩阵形式的高维数据的有力工具. 通过本节可以看出如何将它用于张量场的缩并, 它在物理中, 特别是在相对论的讨论中起着关键性的作用[48, 90].

下面的这段话引自文献[48]: “有时用矩阵方法处理重复下标的和也很方便. 这些方法局限于处理一个或两个下标, 只能允许它们排列成一维序列(行向量或列向量) 或两维序列(矩阵).” 从上面的话中可以看出, 用传统的方法只有1 维或2 维的数据可以看作矩阵形式. 但是用半张量积, 结合张量积和换位矩阵等, 可以很容易将高维数据看成矩阵.

设$\sigma\in T_s^r(V)$ 是n 维向量空间V 上的(r,s) 型张量场, $1\le p\le r$, $1\le q\le s$. 我们定义缩并$\pi_q^p: T_s^r(V)\to T_{s-1}^{r-1}(V)$ 如下: 分别固定V 的基底$\{d_1,\cdots,d_n\}$ 和它对偶空间V^* 的基底$\{e^1,\cdots,e^n\}$, 对于$\omega\in T_s^r(V)$, 记

$$\omega_{j_1\cdots j_s}^{i_1\cdots i_r}=\omega(d_{i_1},\cdots,d_{i_r};e^{j_1},\cdots,e^{j_s}),\quad i_1,\cdots,i_r,j_1,\cdots,j_s=1,\cdots,n.$$

于是我们可以得到一个n^{r+s} 维数据, 将它排列成矩阵

$$M_\omega=\begin{bmatrix}\omega_{1\cdots 11}^{1\cdots 11} & \omega_{1\cdots 11}^{1\cdots 12} & \cdots & \omega_{1\cdots 11}^{n\cdots nn}\\ \omega_{1\cdots 12}^{1\cdots 11} & \omega_{1\cdots 12}^{1\cdots 12} & \cdots & \omega_{1\cdots 12}^{n\cdots nn}\\ \vdots & \vdots & & \vdots\\ \omega_{n\cdots nn}^{1\cdots 11} & \omega_{n\cdots nn}^{1\cdots 12} & \cdots & \omega_{n\cdots nn}^{n\cdots nn}\end{bmatrix}. \tag{7.3.1}$$

即矩阵M_ω 为张量ω 的结构矩阵. 不难验证下面的公式:

$$\begin{aligned}&\omega(\sigma_1,\cdots,\sigma_s;X_1,\cdots,X_r)\\ &\quad=\sigma_s\ltimes\cdots\ltimes\sigma_1\ltimes M_\omega\ltimes X_1\ltimes\cdots\ltimes X_r.\end{aligned} \tag{7.3.2}$$

下面我们用结构矩阵定义缩并$\pi_q^p(\sigma)$, 它的结构矩阵的元素由

$$\omega_{j_1\cdots\hat{j_q}\cdots j_s}^{i_1\cdots\hat{i_p}\cdots i_r} = \sum_{i_p=j_q} \omega_{j_1\cdots j_q\cdots j_s}^{i_1\cdots i_p\cdots i_r} \tag{7.3.3}$$

决定, 这里"ˆ" 表示省略相应的指标.

由于定义依赖于基底的选取(一般对于张量场依赖于坐标的选取), 我们需要证明这个定义与坐标无关. 前面说过, 在下标形式下, 证明很麻烦. 我们将给出一个简洁的矩阵证明. 此外, 我们还将给出缩并的张量的结构矩阵.

首先, 我们给出$\pi_q^p(\sigma)$ 的结构矩阵. 设$\xi = n^{s-1}$, $\eta = n^{r-1}$, 那么结构矩阵M_σ可以写成

$$M_\sigma = \begin{bmatrix} M_{11} & \cdots & M_{1\eta} \\ \vdots & & \vdots \\ M_{\xi 1} & \cdots & M_{\xi\eta} \end{bmatrix}, \tag{7.3.4}$$

其中每个块M_{ij} 是$n \times n$ 矩阵.

直接计算可知

引理 7.3.1 设$p = r$, $q = s$, 则

$$M_{\pi_s^r(\sigma)} = \begin{bmatrix} \mathrm{tr}(M_{11}) & \cdots & \mathrm{tr}(M_{1\eta}) \\ \vdots & & \vdots \\ \mathrm{tr}(M_{\xi 1}) & \cdots & \mathrm{tr}(M_{\xi\eta}) \end{bmatrix} := TR(M_\sigma). \tag{7.3.5}$$

其中算子TR 用于计算所有块的迹.

对于一般情形, 我们需要交换指标p 和r, q 和s. 任意两个元素的交换可以通过相邻两个元素的交换序列实现. 利用命题2.3.19, 我们有交换的结构矩阵

$$\begin{aligned} \tilde{M}_\sigma &= \textstyle\prod_{t=0}^{s-q-1}(I_{n^{s-2-t}} \otimes W_{[n]} \otimes I_{n^t}) M_\sigma \prod_{t=0}^{r-p-1}(I_{n^{r-2-t}} \otimes W_{[n]} \otimes I_{n^t}) \\ &:= \Pi_1 M_\sigma \Pi_2. \end{aligned} \tag{7.3.6}$$

类似于M, 我们可以将$\tilde{M}$ 分割成由$n \times n$ 矩阵组成的$\xi \times \eta$ 个块, 记为$\tilde{M}_{ij}$, 于是我们有

命题 7.3.2 *$\pi_q^p(\sigma)$ 的结构矩阵是*

$$M_{\pi_q^p(\sigma)} = TR(\tilde{M}_\sigma) = TR(\Pi_1 M_\sigma \Pi_2). \tag{7.3.7}$$

我们给出一个例子来说明缩并.

例 7.3.3　设$n=2$, $r=2$, $s=3$. 我们考虑$\pi_1^1(\sigma)$, 记

$$M_\sigma=\begin{bmatrix} a_{111}^{11} & a_{111}^{12} & a_{111}^{21} & a_{111}^{22}\\ a_{112}^{11} & a_{112}^{12} & a_{112}^{21} & a_{112}^{22}\\ a_{121}^{11} & a_{121}^{12} & a_{121}^{21} & a_{121}^{22}\\ a_{122}^{11} & a_{122}^{12} & a_{122}^{21} & a_{122}^{22}\\ a_{211}^{11} & a_{211}^{12} & a_{211}^{21} & a_{211}^{22}\\ a_{212}^{11} & a_{212}^{12} & a_{212}^{21} & a_{212}^{22}\\ a_{221}^{11} & a_{221}^{12} & a_{221}^{21} & a_{221}^{22}\\ a_{222}^{11} & a_{222}^{12} & a_{222}^{21} & a_{222}^{22} \end{bmatrix}.$$

$$\begin{aligned}\Pi_1 &= \prod_{t=0}^{1} I_{2^{1-t}}\otimes W_{[2]}\otimes I_{2^t}=(I_2\otimes W_{[2]})(W_{[2]}\otimes I_2)\\ &=\begin{bmatrix}1&0&0&0&0&0&0&0\\0&0&0&0&1&0&0&0\\0&1&0&0&0&0&0&0\\0&0&0&0&0&1&0&0\\0&0&1&0&0&0&0&0\\0&0&0&0&0&0&1&0\\0&0&0&1&0&0&0&0\\0&0&0&0&0&0&0&1\end{bmatrix}.\end{aligned}$$

$$\Pi_2=W_{[2]}=\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}.$$

于是

$$\Pi_1M_\sigma\Pi_2=\begin{bmatrix} a_{111}^{11} & a_{111}^{21} & a_{111}^{12} & a_{111}^{22}\\ a_{211}^{11} & a_{211}^{21} & a_{211}^{12} & a_{211}^{22}\\ a_{112}^{11} & a_{112}^{21} & a_{112}^{12} & a_{112}^{22}\\ a_{212}^{11} & a_{212}^{21} & a_{212}^{12} & a_{212}^{22}\\ a_{121}^{11} & a_{121}^{21} & a_{121}^{12} & a_{121}^{22}\\ a_{221}^{11} & a_{221}^{21} & a_{221}^{12} & a_{221}^{22}\\ a_{122}^{11} & a_{122}^{21} & a_{122}^{12} & a_{122}^{22}\\ a_{222}^{11} & a_{222}^{21} & a_{222}^{12} & a_{222}^{22} \end{bmatrix}.$$

利用上式和命题7.3.2 得到

$$M_{\pi_1^1(\sigma)} = \begin{bmatrix} a_{111}^{11} + a_{211}^{21} & a_{111}^{12} + a_{211}^{22} \\ a_{112}^{11} + a_{212}^{21} & a_{112}^{12} + a_{212}^{22} \\ a_{121}^{11} + a_{221}^{21} & a_{121}^{12} + a_{221}^{22} \\ a_{122}^{11} + a_{222}^{21} & a_{122}^{12} + a_{222}^{22} \end{bmatrix}.$$

□

接下来, 我们证明(7.3.6) 所定义的结构矩阵与坐标无关. 假设我们有一个坐标变换$z = z(x)$, 并且它的Jacobi 矩阵是$J = \frac{\partial z}{\partial x}$.

通过直接计算可以验证下面的引理.

引理 7.3.4 (1) 设$P \in M_{s\times m}$, $Q \in M_{n\times n}$, $A_i \in M_{n\times rn}$, $i = 1, \cdots, m$. 令

$$\tilde{A} = \begin{bmatrix} \tilde{A}_1 \\ \vdots \\ \tilde{A}_m \end{bmatrix} = (P \otimes Q) \begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} = (P \otimes Q)A.$$

则

$$\begin{bmatrix} \tilde{A}_1 \\ \vdots \\ \tilde{A}_m \end{bmatrix} = P \ltimes \begin{bmatrix} QA_1 \\ \vdots \\ QA_m \end{bmatrix},$$

且

$$TR(\tilde{A}) = P \cdot TR\left(\begin{bmatrix} QA_1 \\ \vdots \\ QA_m \end{bmatrix}\right).$$

(2) 设$P \in M_{m\times s}$, $Q \in M_{n\times n}$, $A_i \in M_{nr\times n}$, $i = 1, \cdots, m$. 令

$$\tilde{A} = \left[\tilde{A}_1, \cdots, \tilde{A}_m\right] = [A_1, \cdots, A_m]\,(P \otimes Q) = A(P \otimes Q).$$

则

$$\left[\tilde{A}_1, \cdots, \tilde{A}_m\right] = [A_1Q, \cdots, A_mQ]\,P,$$

且

$$TR(\tilde{A}) = TR\,(A_1Q, \cdots, A_mQ)\,P.$$

现在准备证明本节的主要结果: 缩并是定义好的. 显然只要证明它与局部坐标无关, 则它是定义好的.

定理 7.3.5 (7.3.3) 所定义的缩并与坐标无关.

证明 设

$$M_\sigma = \begin{bmatrix} M_{11} & \cdots & M_{1\eta} \\ \vdots & & \vdots \\ M_{\xi 1} & \cdots & M_{\xi\eta} \end{bmatrix}.$$

利用命题7.3.2, 我们有

$$M_{\pi_q^p(\sigma)} = TR(\Pi_1 M_\sigma \Pi_2).$$

现在考虑坐标变换$z = z(x)$, 于是M_σ 变为

$$\bar{M}_\sigma = \underbrace{J^{-1} \otimes \cdots \otimes J^{-1}}_{s} M_\sigma \underbrace{J \otimes \cdots \otimes J}_{t}.$$

注意到Π_1 与$\underbrace{J^{-1} \otimes \cdots \otimes J^{-1}}_{s}$ 是可交换的, Π_2 与$\underbrace{J \otimes \cdots \otimes J}_{t}$ 是可交换的. 对$\bar{M}_\sigma$ 应用命题7.3.2, 我们有

$$\begin{aligned}
\bar{M}_{\pi_q^p(\sigma)} &= TR\left(\Pi_1 \underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s} M_\sigma \underbrace{(J \otimes \cdots \otimes J)}_{t} \Pi_2\right) \\
&= TR\left(\underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s} (\Pi_1 M_\sigma \Pi_2) \underbrace{(J \otimes \cdots \otimes J)}_{t}\right) \\
&= TR\left((\underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s-1} \otimes J^{-1})(\tilde{M}_\sigma)(\underbrace{(J \otimes \cdots \otimes J)}_{t-1} \otimes J)\right) \\
&= \underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s-1} TR(J^{-1}(\tilde{M}_\sigma)J) \underbrace{(J \otimes \cdots \otimes J)}_{t-1} \\
&= \underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s-1} TR(\tilde{M}_\sigma) \underbrace{(J \otimes \cdots \otimes J)}_{t-1} \\
&= \underbrace{(J^{-1} \otimes \cdots \otimes J^{-1})}_{s-1} M_{\pi_q^p(\sigma)} \underbrace{(J \otimes \cdots \otimes J)}_{t-1}.
\end{aligned}$$

注意最后三个等式可由引理7.3.4 得到. □

7.4 注释与参考

文献[10] 是一本很好的关于微分流形及黎曼几何的入门书. 它的最大特点是注重实用和可计算性. 本章关于联络、曲率算子、黎曼曲率张量等概念、记号及基本计算公式均采自此书.

关于代数的概念, 在第5 章已有所讨论. 欲了解更多的详情可参见标准近世代数教程, 例如文献[67]. 关于李代数的书也很多, 例如文献[66] 就是一本不错的参考书.

四元数在力学系统控制中有许多应用, 用它刻画姿态角可避免奇异性, 参见文献[76].

关于张量分析及张量场, 可参见文献[46], [10]. 关于张量缩并及其在相对论中的应用,可参见文献[48].

本章内容主要基于以下几篇文章: 第1 节基于文献[21], 第2 节基于文献[28], 第3 节基于文献[33].

第 8 章　非线性控制系统的镇定

一般地说, 一个动力系统要能够正常工作, 系统的稳定性是一个基本的要求. 因此, 控制系统的一个重要任务就是设计一个控制器使系统稳定, 这样的问题称为镇定问题.

动力系统的中心流形理论是20世纪70年代前后发展起来的. 此后, 它被广泛应用于非线性系统的镇定设计中. 由于在中心流形设计中, 多项式近似起着重要作用. 因此, 多元多项式的半张量积表示成为一个有效的工具. 本章介绍如何通过反馈控制来设计中心流形, 从而镇定系统, 以及如何在中心流形设计中使用半张量积方法.

8.1　非线性控制系统

非线性控制系统是本章的研究对象, 我们主要讨论一类有广泛应用背景的非线性系统, 称为仿射非线性控制系统, "仿射" 指它对控制是线性的. 一个仿射非线性控制系统可表示为

$$\dot{\xi} = f(\xi) + \sum_{i=1}^{m} g_i(\xi)u_i, \quad \xi \in \mathbb{R}^{\ell}, \tag{8.1.1}$$

这里$u = (u_1, \cdots, u_m)$ 为控制, $f(0) = 0$, 并且, 为讨论方便, 我们设$f(\xi)$ 及$g_i(\xi)$ 均为解析向量场. 本节介绍一些非线性控制的基本结果而不加证明, 不熟悉非线性控制的读者可参阅文献[68] 或[71] 中的相关内容.

在一定的正则性条件下, 系统(8.1.1) 可经过坐标变换和状态反馈变为下面的标准形式, 称为Byrnes-Isidori 正则型

$$\begin{cases} \dot{x} = Ax + Bu, \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m; \\ \dot{z} = q(x, z), \quad z \in \mathbb{R}^s. \end{cases} \tag{8.1.2}$$

这里$n + s = \ell$, (A, B) 是完全能控对. 考虑它的非线性部分, 称下面的系统为原系统(8.1.1) 或(8.1.2) 的零动态:

$$\dot{z} = q(0, z). \tag{8.1.3}$$

如果系统(8.1.3) 在原点是渐近稳定的, 则称原系统(8.1.2) 具有最小相位, 否则, 称原系统具有非最小相位. 这些名词均来自相应的线性系统, 而且它们与系统输出有关. 为方便, 这里的说法直观、明确, 但从控制理论看不够严格.

本章讨论的镇定问题指, 如可能, 找出状态反馈控制

$$u = u(x, z),$$

使得闭环系统

$$\begin{cases} \dot{x} = Ax + Bu(x, z), \\ \dot{z} = q(x, z) \end{cases} \tag{8.1.4}$$

在原点渐近稳定.

下面的引理[68]对最小相位系统的镇定十分重要.

引理 8.1.1 *考虑系统*

$$\begin{cases} \dot{z} = Az + p(z, w), \\ \dot{w} = f(z, w). \end{cases} \tag{8.1.5}$$

这里, 设对所有位于0 附近的w 均成立$p(0, w) = 0$, 并且

$$\frac{\partial p}{\partial z}(0, 0) = 0.$$

如果$w = 0$ 是$\dot{w} = f(0, w)$ 的渐近稳定平衡点, 并且A 是Hurwitz 矩阵, 那么$(z, w) = (0, 0)$ 是系统(8.1.5) 的渐近稳定平衡点.

比较(8.1.5) 与(8.1.2), 一个线性反馈控制

$$u = Kx$$

能够镇定整个系统, 只要选择合适的K 使$A + BK$ 为Hurwitz 矩阵即可.

因此, 我们更感兴趣的情形, 也是更具挑战性的情形, 是系统具有非最小相位的情形.

为了后面讨论的方便, 我们给出Byrnes-Isidori 正则型一个更具体的表现形式. 因为在Byrnes-Isidori 正则型中(A, B) 是一个完全能控对, 我们可以将这一部分写成Brunovsky 线性能控标准形式, 即

$$\begin{cases} \dot{x}_1^i = x_2^i, \\ \vdots \\ \dot{x}_{n_i-1}^i = x_{n_i}^i, \\ \dot{x}_{n_i}^i = u_i, \\ \qquad\qquad i = 1, \cdots, m, \\ \dot{z} = p(x, z), z \in \mathbb{R}^s. \end{cases} \tag{8.1.6}$$

这里$\sum_{i=1}^{m} n_i = n$, $d_i(z,\xi) = (d_{i1}(x,z), \cdots, d_{im}(x,z))$, $i = 1, \cdots, m$.

下面我们将集中讨论这个模型.

8.2 中心流形理论

本节给出中心流形理论的一些相关的基本知识, 证明及相关的细节可参考文献[14] 和[71].

考虑一个动态系统

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n, \tag{8.2.1}$$

这里$f(x)$ 是一个C^ω 向量场. 下面是中心流形理论的一个基本定理.

定理 8.2.1 设x_e 为系统(8.2.1) 的一个平衡点. 则存在三个子流形S^+, S^- 和S^0, 使得

(1) 它们均为通过x_e 的, 关于系统(8.2.1) 的不变子流形;

(2) 它们相应的切空间$T_{x_e}(S^+)$, $T_{x_e}(S^-)$ 和$T_{x_e}(S^0)$ 线性无关, 并且

$$T_{x_e}(S^+) \oplus T_{x_e}(S^-) \oplus T_{x_e}(S^0) = T_{x_e}(\mathbb{R}^n);$$

(3) 将系统(8.2.1) 限制于S^+ (相应地, S^-, S^0), 那么,它在S^+ (相应地, S^- 或S^0) 上的动力系统在x_e 点的线性近似的特征根均具有正实部(相应地, 负实部或零实部);

(4) S^+ 及S^- 均唯一, 但是S^0 可能不唯一.

S^+, S^- 和S^0 分别称为稳定子流形, 不稳定子流形和中心流形.

中心流形是一个局部结构, 我们只在x_e 的一个邻域U 上来讨论它. 因此, 本章讨论的镇定问题均指局部镇定.

下面用一个简单例子来刻画这几个不变子流形:

例 8.2.2 考虑下述系统

$$\begin{cases} \dot{x}_1 = 2x_2 - \sin(x_3)(x_4 + x_3^2), \\ \dot{x}_2 = -2x_1 - 4x_2 + x_1 x_3, \\ \dot{x}_3 = x_3, \\ \dot{x}_4 = x_4^2 + 2(x_4 - 1)x_3^2 + x_3^4. \end{cases} \tag{8.2.2}$$

显然, 0 是它的一个平衡点, 我们考虑这个平衡点的三个不变子流形S^+, S^-和S^0. 可以证明, 它们分别是

$$\begin{aligned} S^+ &= \{x\in\mathbb{R}^4 \mid x_1=0, x_2=0, x_4+x_3^2=0\};\\ S^- &= \{x\in\mathbb{R}^4 \mid x_3=0, x_4=0\};\\ S^0 &= \{x\in\mathbb{R}^4 \mid x_1=0, x_2=0, x_3=0\}. \end{aligned}$$

先看S^+: 首先我们证明S^+ 是不变流形. 从S^+ 上的点出发, 我们有$x_1(0)=0$, $x_2(0)=0$, 并且

$$\begin{cases} \dot{x}_1 = 0,\\ \dot{x}_2 = 0. \end{cases}$$

即$x_1(t)=0$, $x_2(t)=0$, $\forall t>0$. 令$\xi = x_4+x_3^2$, 同样, 从S^+ 上的点出发, 我们有$\xi(0)=0$. 并且,

$$\dot{\xi} = \xi^2.$$

由此可知$\xi(t)=0$, $\forall t>0$. 综上可得S^+ 是不变子流形. 然后考虑系统(8.2.2) 限制在S^+ 上的子系统. 它是1 维子系统, 我们不妨以x_3 为其坐标. 于是可得其动态方程为

$$\dot{x}_3 = x_3.$$

显然它是不稳定子流形. 值得说明的是, 坐标的选择不影响其上动态系统的动力学本质, 例如, 我们也可以选x_4 作为该流形的坐标, 则得

$$\dot{x}_4 = x_4^2 + 2(x_4-1)x_3^2 + x_3^4 = x_4^2 + 2(x_4-1)(-x_4) + (-x_4^2) = 2x_4.$$

故结论不变. 类此, 我们可以检验S^- 是不变流形, 其上的动力系统为

$$\begin{cases} \dot{x}_1 = 2x_2,\\ \dot{x}_2 = -2x_1-4x_2. \end{cases}$$

它是稳定子流形; S^0 是不变流形, 其上的动力系统为

$$\dot{x}_4 = x_4^2,$$

它的线性部分为零. □

下面的三个定理是中心流形理论的主要结果.

定理 8.2.3 考虑非线性系统

$$\begin{cases} \dot{x} = Ax + p(x,z), & x \in \mathbb{R}^n, \\ \dot{z} = Cz + q(x,z), & z \in \mathbb{R}^m, \end{cases} \tag{8.2.3}$$

这里, $\mathrm{Re}\,\sigma(A) < 0$, $\mathrm{Re}\,\sigma(C) = 0$, $p(x,z)$ 和$q(x,z)$ 及其一阶导数在原点均为零. 则存在一个过原点的m 维不变流形(即中心流形), 它可由下述方程描述:

$$S = \{(x,z) \mid x = h(z)\}, \tag{8.2.4}$$

这里$h(z)$ 满足

$$\frac{\partial h(z)}{\partial z}(Cz + q(h(z),z)) - Ah(z) - p(h(z),z) = 0. \tag{8.2.5}$$

定理 8.2.4 (1) 中心流形上的动态方程为

$$\dot{z} = Cz + q(h(z),z), \quad z \in \mathbb{R}^m. \tag{8.2.6}$$

(2) 系统(8.2.3) 在原点是渐近稳定、稳定或不稳定的, 当且仅当, 中心流形上的动力系统(8.2.6) 是渐近稳定,稳定或不稳定的.

定理 8.2.5 如果存在一个光滑函数$\phi(z)$, 使得下式定义的算子M 满足

$$M(\phi(z)) := \frac{\partial \phi(z)}{\partial z}(Cz + q(\phi(z),z)) - A\phi(z) - p(\phi(z),z) = O(\|z\|^{k+1}). \tag{8.2.7}$$

那么它与中心流形方程之差满足

$$\|\phi(z) - h(z)\| = O(\|z\|^{k+1}). \tag{8.2.8}$$

为直观和方便计, 我们分别将定理8.2.3, 8.2.4 和8.2.5 称为中心流形的存在定理, 等价定理和逼近定理.

8.3 镇定与导数齐次Lyapunov 函数

通过状态反馈设计闭环系统的中心流形, 从而达到镇定系统的目的, 这是本章的基本思路. 下面, 我们先用一个简单例子来说明这个设计思想.

例 8.3.1 考虑下述系统

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = x_3, \\ \dot{x}_3 = f(x,z) + g(x,z)u, & g(0,0) \neq 0, \\ \dot{z} = z^3 + zx_1. \end{cases} \tag{8.3.1}$$

系统(8.3.1) 已经是一个标准形式, 它的零动态为

$$\dot{z} = z^3, \tag{8.3.2}$$

它在原点不稳定, 故(8.3.1) 是一个非最小相位系统. 因此, 准线性控制

$$u = -\frac{f(x,z)}{g(x,z)} + \frac{1}{g(x,z)}(a_1x_1 + a_2x_2 + a_3x_3)$$

不能镇定原系统. 我们须考虑非线性反馈.

我们不妨考虑如下控制

$$u = -\frac{f(x,z)}{g(x,z)} + \frac{1}{g(x,z)}(a_1x_1 + a_2x_2 + a_3x_3 + bz^2). \tag{8.3.3}$$

为镇定系统, 先选a_1, a_2, a_3 以镇定能控的线性部分的状态变量x_1, x_2, x_3. 然后, 选择b 来镇定中心流形变量z. 为求得可能的b 值, 我们考虑中心流形方法. 尝试用

$$\phi(z) = \begin{bmatrix} \phi_1(z) \\ \phi_2(z) \\ \phi_3(z) \end{bmatrix} = O(\|z\|^2)$$

来逼近中心流形. 根据定理8.2.5, 对系统(8.3.1) 我们有

$$\begin{aligned} M\phi(z) &= \frac{\partial \phi(z)}{\partial z}(Cz + q(\phi(z), z)) - A\phi(z) - p(\phi(z), z) \\ &= D\phi(z)\,(q(z)\phi_1(z)) - \begin{bmatrix} \phi_2(z) \\ \phi_3(z) \\ a_1\phi_1(z) + a_2\phi_2(z) + a_3\phi_3(z) + bz^2 \end{bmatrix} \\ &= O(\|z\|^4) - \begin{bmatrix} \phi_2(z) \\ \phi_3(z) \\ a_1\phi_1(z) + a_2\phi_2(z) + a_3\phi_3(z) + bz^2 \end{bmatrix}. \end{aligned}$$

选择

$$\begin{cases} \phi_1 = -\dfrac{b}{a_1}z^2, \\ \phi_i = 0, \quad i = 2, 3. \end{cases}$$

则可得$M\phi(z) = O(\|z\|^4)$. 根据逼近定理, 可知中心流形可表示为

$$\begin{cases} x_1 = h_1(z) = -\dfrac{b}{a_1}z^2 + O(\|z\|^4), \\ x_i = h_i(z) = O(\|z\|^4), \quad i = 2, 3. \end{cases} \tag{8.3.4}$$

中心流形上的动态系统为

$$\dot{z} = \left(1 - \frac{b}{a_1}\right) z^3 + O(\|z\|^4). \tag{8.3.5}$$

最后, 选择$\{a_1, a_2, a_3, b\}$ 使得线性部分为Hurwitz 矩阵, 即稳定矩阵, 并且

$$1 - \frac{b}{a_1} < 0,$$

例如, 选$a_1 = -1$, $a_2 = a_3 = -3$, $b = -2$. 则反馈控制变为

$$u = -\frac{f(x,z)}{g(x,z)} + \frac{1}{g(x,z)}(-x_1 - 3x_2 - 3x_3 - 2z^2).$$

于是, 系统(8.3.5) 在原点渐近稳定. 再根据等价定理可知, 闭环系统在原点也是渐近稳定的. 即, 上述控制镇定了整个系统. □

由上述例子可知

$$u = -\frac{f(x)}{g(x)} + \frac{1}{g(x)}\left[Kx + P(z)\right]$$

可以用来镇定控制系统, 这里, 线性部分Kx 用来镇定线性能控子系统, 即, 选择K 使$A + BK$ 为一Hurwitz 矩阵. $P(z)$ 为z 的一个多项式, $P(z) = O(\|z\|^2)$, 它用来设计一个稳定的中心流形.

因为中心流形的表达式$h(z)$ 满足一个偏微分方程, 所以一般是无法求出的. 解决的办法是, 用中心流形逼近定理求它的近似方程, 再由近似方程的稳定性判断原系统的稳定性. 因此, 如何由近似方程稳定推出原动态系统的稳定性成为一个关键问题. 导数齐次Lyapunov 函数就是为了这一目的而提出的新工具. 下面我们介绍这一工具.

考虑动力系统

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n, \tag{8.3.6}$$

这里$f(0) = 0$.

我们对f 的每个分量f_i 作Taylor 展开如下

$$f_i(x) = C^i_{k_i} x^{k_i} + O(\|x\|^{k_i+1}), \quad i = 1, \cdots, n, \tag{8.3.7}$$

这里$C^i_{k_i} \neq 0$, 即k_i 是f_i 的最低非零项指数.

然后, 我们给出关于近似系统的以下定义:

定义 8.3.2 (1) 设如(8.3.7), k_i 为$f_i(x)$, $i=1,\cdots,n$ Taylor 展开后最低阶非零项的指数. 定义

$$\dot{x}=g(x),\quad x\in\mathbb{R}^n, \tag{8.3.8}$$

这里$g(x)=(g_1(x),\cdots,g_n(x))^{\mathrm{T}}$ 且

$$g_i=C_{k_i}^i x^{k_i},\quad i=1,\cdots,n.$$

系统(8.3.8) 称为系统(8.3.6) 的近似系统.

(2) 系统(8.3.8) 称为系统(8.3.6) 的奇近似系统, 如果所有的k_i 均为奇数.

(3) 系统(8.3.6) 称为依指数$(k_1,\cdots,k_n)$ 近似稳定的, 如果

$$\dot{x}_i=f_i(x)+O(\|x\|^{k_i+1}),\quad i=1,\cdots,n$$

在原点渐近稳定.

注 (1) 在(8.3.8) 中g_i 是k_i 次齐次多项式. 因此, $g=(g_1,\cdots,g_n)^{\mathrm{T}}$ 称为分量齐次的多项式向量场.

(2) 如果$k_1=\cdots=k_n:=k$, 那么, 这里定义的近似稳定就与经典的近似稳定一致[61]. 在一般情况下, 它与坐标选择有关. 显然, 近似稳定蕴含渐近稳定, 但渐近稳定推不出近似稳定.

定义 8.3.3 给定一个分量齐次的多项式向量场$g=(g_1,\cdots,g_n)^{\mathrm{T}}$, 一个正定多项式$V>0$ 称为g 的导数齐次 Lyapunov 函数 (LFHD), 如果它对于g 的李导数L_gV 是齐次多项式, 即

$$\deg(L_gV)=\deg\left(\frac{\partial V}{\partial x_i}\right)+\deg(g_i),\quad \forall i=1,\cdots,n.$$

下面的例子给出两个典型的导数齐次Lyapunov 函数的构造方法.

例 8.3.4 设$g=(g_1,\cdots,g_n)^{\mathrm{T}}$ 为一个分量齐次的多项式向量场, 且$\deg(g_i)=k_i$, $i=1,\cdots,n$, 设m 为一给定正整数, 满足

$$2m\geq\max\{k_1,\cdots,k_n\}+1.$$

(1) 设$2m_i=2m-k_i+1$, $i=1,\cdots,n$, 那么

$$V=\sum_{i=1}^n p_i x_i^{2m_i} \tag{8.3.9}$$

是g 的一个导数齐次Lyapunov 函数, 只要$p_i > 0$, $\forall i$.

(2) 设 $k_1 = \cdots = k_{n_1} := k^1; k_{n_1+1} = \cdots = k_{n_1+n_2} := k^2; \cdots; k_{n_1+\cdots+n_{r-1}+1} = \cdots = k_{n_1+\cdots+n_r} := k^r$, 这里$k^i$ 为奇数, 且$\sum\limits_{i=1}^{r} n_i = n$. 记$x = (x^1, \cdots, x^r)$, 这里$\dim(x^i) = n_i$, 再设$2m^i = 2m - k^i + 1$, $i = 1, \cdots, r$, 那么

$$V = \sum_{i=1}^{r} \left((x_1^i)^{m^i}, \cdots, (x_{n_i}^i)^{m^i}\right) P_i \left((x_1^i)^{m^i}, \cdots, (x_{n_i}^i)^{m^i}\right)^{\mathrm{T}} \tag{8.3.10}$$

是g 的一个导数齐次Lyapunov 函数, 只要P_i, $i = 1, \cdots, r$ 是维数为$n_i \times n_i$ 的正定矩阵. □

注意, 不管在(8.3.9) 或(8.3.10) 中, V 对g 的李导数均为阶数为$2m$ 的齐次多项式.

下面的命题是导数齐次Lyapunov 函数的一个基本性质.

命题 8.3.5 *考虑系统(8.3.6), 如果对它的近似系统(8.3.8) 存在一个导数齐次Lyapunov 函数, 使其沿(8.3.8) 的导数是负定的, 那么系统(8.3.6) 在原点近似稳定.*

证明 假定L_gV 负定, 并且它的指数为偶数, 设$\deg(L_gV) = 2m$. 我们断言, 存在一个实数$b > 0$ 使得

$$L_gV(x) \le -b\sum_{i=1}^{n}(x_i)^{2m}. \tag{8.3.11}$$

由于L_gV 是负定的, 在紧球面

$$S = \left\{ z \;\middle|\; \sum_{i=1}^{n}(z_i)^{2m} = 1 \right\}$$

上$L_gV(x)$ 达到它的最大值$-b < 0$. 即

$$L_gV(z) \le -b < 0, \quad z \in S.$$

现在, 任何一个$x \in \mathbb{R}^n$ 均可表示为$x = kz$, 这里$z \in S$. 因此

$$L_gV(x) = L_gV(kz) = k^{2m}L_gV(z) \le -bk^{2m} = -b\sum_{i=1}^{n}(x_i)^{2m},$$

这就证明了我们的断言.

利用(8.3.11), 该导数齐次Lyapunov 函数的导数为

$$\dot{V}|_f = \dot{V}|_{g+O(\|x\|^{K+1})} \le -b\sum_{i=1}^{n}(x_i)^{2m} + O(\|x\|^{2m+1}), \tag{8.3.12}$$

这里$g + O(\|x\|^{K+1})$ 是

$$\tilde{g}(x) = \left(g_1(x) + O(\|x\|^{k_1+1}), \cdots, g_n(x) + O(\|x\|^{k_n+1})\right)^{\mathrm{T}}.$$

的缩写形式. □

对于齐次向量场我们有如下近似稳定性的结果[57] (只是在叙述上略有不同):

定理 8.3.6 设系统(8.3.6) 的具有指数$k_1 = \cdots = k_n = k$ 的近似系统(8.3.8) 渐近稳定, 则系统(8.3.6) 渐近稳定.

命题8.3.5 和定理8.3.6 是我们判定近似稳定的主要工具.

8.4 齐次多项式的负定性

在导数齐次Lyapunov 函数方法中, 一个关键问题是判定其导数的负定性. 由于其导数是一个齐次多项式, 故要判定一个齐次多项式的负定性. 本节的目的是要发展出一套实用的判别方法.

下面的不等式是我们的出发点.

引理 8.4.1 设$S \in \mathbb{Z}_+^n$ 及$x \in \mathbb{R}^n$. 那么,下面的不等式成立:

$$|x^S| \le \sum_{j=1}^{n} \frac{S_j}{|S|}|x_j|^{|S|}. \tag{8.4.1}$$

证明 设$z_1, \cdots, z_n \ge 0$ 及$s_1, \cdots, s_n \ge 0$. 由于$\ln(z)$ 是一个凸函数, 我们有

$$\sum_{i=1}^{n}\left(\frac{s_i}{\sum\limits_{k=1}^{n} s_k}\ln(z_i)\right) \le \ln\left(\sum_{i=1}^{n}\frac{s_i}{\sum\limits_{k=1}^{n} s_k} z_i\right).$$

记$s = \sum\limits_{k=1}^{n} s_k$, 则上式可等价地表示成

$$\ln\prod_{i=1}^{n} z_i^{\left(\frac{s_i}{s}\right)} \le \ln\left(\sum_{i=1}^{n}\frac{s_i}{s}z_i\right),$$

或

$$\prod_{i=1}^{n} z_i^{\left(\frac{s_i}{s}\right)} \le \sum_{i=1}^{n} \frac{s_i}{s} z_i.$$

将$z_i^{\left(\frac{1}{s}\right)}$ 用$|x_i|$ 代入, 立即得到不等式(8.4.1). □

为使用上述引理, 我们必须知道对于n 元k 次齐次式x^k, $x \in \mathbb{R}^n$, 每个变元x_i 在其某个分量中的指数. 因为x^k 有n^k 个分量. 于是对每一个变量x_i, 我们可以用一个n^k 维向量, 记作V_k^i 来表示x^k 各分量中x_i 的指数.

例 8.4.2 设$x \in \mathbb{R}^2$, 可算出

$$\begin{aligned} x^4 \;=\; & (x_1^4, x_1^3x_2, x_1^2x_2x_1, x_1^2x_2^2, x_1x_2x_1^2, x_1x_2x_1x_2, x_1x_2^2x_1, x_1x_2^3, \\ & x_2x_1^3, x_2x_1^2x_2, x_2x_1x_2x_1, x_2x_1x_2^2, x_2^2x_1^2, x_2^2x_1x_2, x_2^3x_1, x_2^4). \end{aligned}$$

于是有

$$V_4^1 = [4,3,3,2,3,2,2,1,3,2,2,1,2,1,1,0]^{\mathrm{T}},$$

及

$$V_4^2 = [0,1,1,2,1,2,2,3,1,2,2,3,2,3,3,4]^{\mathrm{T}}.$$

□

下面讨论V_k^i 的一般表达式.

引理 8.4.3 *设$x \in \mathbb{R}^n$, 变量x_i 在x^k 各分量中的指数, 用向量V_k^i 表示, 则*

$$V_k^i = \mathbf{1}_n^{k-1}\delta_i^n + \mathbf{1}_n^{k-2}\delta_i^n\mathbf{1}_n^1 + \cdots + \mathbf{1}_n\delta_i^n\mathbf{1}_n^{k-2} + \delta_i^n\mathbf{1}_n^{k-1}, \tag{8.4.2}$$

这里$\mathbf{1}_n = (\underbrace{1,\cdots,1}_{n})^{\mathrm{T}}$, δ_i^n 为单位阵I_n 的第i 列.

证明 先证明一个递推公式

$$\begin{cases} V_1^i = \delta_i^n, \\ V_{s+1}^i = \mathbf{1}_n V_s^i + \delta_i^n \mathbf{1}_{n^{s-1}}, \quad s \ge 1. \end{cases} \tag{8.4.3}$$

由于$x^1 = (x_1, x_2, \cdots, x_n)^{\mathrm{T}}$, 故$x_i$ 只在其第i 个分量出现一次, 因此, $V_1^i = \delta_i^n$. 现在假定V_s^i 已知, 它是一个n^s 维向量, 要从V_s^i 得到V_{s+1}^i, 可以通过以下两个步骤实现: 首先, 将其重复排n 次, 得到一个n^{s+1} 维向量. 它是由于x^s 要分别与$x_1, \cdots, x_n$ 相乘而得到的. 这是由$\mathbf{1}_n V_s^i$ 来实现. 其次, x^{s+1} 的第i 个n^s 维块是由x^s 与x_i 相

乘而得到的. 因此, 在这一块要将x_i 的指数加1, 这可由$\delta_i^n \mathbf{1}_n$ 实现. 综合这两个步骤, 即得(8.4.3).

反复利用(8.4.3) 叠代, 即可得到(8.4.2). □

在x^k 中, x_i 的最高次方项x_i^k 特别重要, 当$k=2$ 时, 这种项称为对角元. 因为它对应于二次型矩阵表示$x^{\mathrm{T}}Qx$ 中Q 的对角元素. 关于二次型负定的判定, 我们有如下的对角占优原理: 二次型$x^{\mathrm{T}}Qx$ 负定, 如果

$$-q_{ii} > \sum_{j\neq i} q_{ij}, \quad i=1,\cdots,n. \tag{8.4.4}$$

对于$k>2$ 的情况, 我们仍将其最高次方项x_i^k 称为对角元. 我们将在本节将对角占优原理推广到高次方的情况. 首先, 我们要知道对角元在x^k 中的位置. 容易证明以下结论:

引理 8.4.4 *设$x\in\mathbb{R}^n$. 在x^k 中对角元x_i^k, $i=1,\cdots,n$ 的位置为*

$$d_i=(i-1)\frac{n^k-1}{n-1}+1, \quad i=1,2,\cdots,n.$$

例如, 设$x\in\mathbb{R}^4$ 且$k=2$. 利用上式可算出: $d_1=1$, $d_2=6$, $d_3=11$, $d_4=16$. 回到例8.4.2, 容易检验, d_i 是x_i^2 所在位置.

为方便计, 我们定义一个对角元位置集合$D_n^k=\{d_i\,|\,i=1,\cdots,n\}$, 这里$d_i$ 是x_i^k 在x^k 中的位置.

利用引理8.4.1, 8.4.3 和8.4.4, 我们有如下的不等式估计

引理 8.4.5 *设k 为一偶数, $P(x)=Fx^k$ 为一k 次齐次多项式. 那么, 我们有如下的不等式估计:*

$$P(x)\le \frac{1}{k}\tilde{F}\left(V_k^1 x_1^k+\cdots+V_k^n x_n^k\right), \tag{8.4.5}$$

这里$\tilde{F}\in\mathbb{R}^{n^k}$ 为一行向量, 其元素定义为

$$\tilde{F}_i=\begin{cases} F_i, & i\in D_n^k;\\ |F_i|, & 其他. \end{cases}$$

证明 将不等式(8.4.1) 用于$P(x)$ 的每一项, 对x^k 的每一个分量, 其x_i^k 系数为$\frac{V_k^i}{k}$. 我们保持对角元x_i^k, $i=1,\cdots,n$ 不变, 并按不等式(8.4.1) 将所有其他项放大, 则得(8.4.5). □

我们用一个例子说明如何用上述方法检验多项式的负定性. 注意, 一个多元多项式如果最低次项负定则在原点附近局部负定.

例 8.4.6　(1) 考虑多项式

$$P(x) = -x_1^4 + x_1^2x_2^2 - 2x_1x_3^3 - 2x_2^4.$$

将其表示为$P(x) = Fx^4$, 则

$$F = [-1, 0, 0, 1, 0, 0, 0, -2, 0, 0, 0, 0, 0, 0, 0, -3],$$

利用引理8.4.5 构造$\tilde{F}$ 如下

$$\tilde{F} = [-1, 0, 0, 1, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, -3].$$

同时, 不难算出

$$\begin{aligned} V_4^1 &= \mathbf{1}_2^3\delta_2^1 + \mathbf{1}_2^2\delta_2^1\mathbf{1}_2 + \mathbf{1}_2\delta_2^1\mathbf{1}_2^2 + \delta_2^1\mathbf{1}_2^3 \\ &= [4, 3, 3, 2, 3, 2, 2, 1, 3, 2, 2, 1, 2, 1, 1, 0]^{\mathrm{T}}, \end{aligned}$$

和

$$\begin{aligned} V_4^2 &= \mathbf{1}_2^3\delta_2^2 + \mathbf{1}_2^2\delta_2^2\mathbf{1}_2 + \mathbf{1}_2\delta_2^2\mathbf{1}_2^2 + \delta_2^2\mathbf{1}_2^3 \\ &= [0, 1, 1, 2, 1, 2, 2, 3, 1, 2, 2, 3, 2, 3, 3, 4]^{\mathrm{T}}, \end{aligned}$$

根据(8.4.5) 有

$$\begin{aligned} P(x) &\le \tfrac{1}{4}\left[\tilde{F}V_4^1 + \tilde{F}V_4^2\right] \\ &= \tfrac{-4+2+1}{4}x_1^4 + \tfrac{2+2\times3-12}{4}x_2^4 \\ &= -\tfrac{1}{4}x_1^4 + x_2^4. \end{aligned}$$

因此, $P(x) < 0$, 即$P(x)$ 为正定多项式.

(2) 考虑

$$Q(x) = -x_1^4 - 6x_1^2x_2^2 - 2x_1x_3^3 - 2x_2^4.$$

将其表示为$Q(x) = Hx^4$, 则

$$H = [-1, 0, 0, -6, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, -3],$$

构造相应的$\tilde{H}$ 为

$$\tilde{H} = [-1, 0, 0, 6, 0, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, -3].$$

利用(8.4.5) 有

$$\begin{aligned} Q(x) &\le \tfrac{1}{4}\left[\tilde{H}V_4^1+\tilde{H}V_4^2\right] \\ &= \tfrac{-4+12+1}{4}x_1^4+\tfrac{12+2\times3-12}{4}x_2^4 \\ &= \tfrac{9}{4}x_1^4+\tfrac{3}{2}x_2^4. \end{aligned}$$

我们从这里得不到任何结论. □

比较上例中的$P(x)$ 和$Q(x)$, 容易看出$P(x)\le Q(x)$, 因此, 不等式(8.4.5) 在这种情况下明显不够精确. 问题出在, 我们事实上不需要对$Q(x)$ 中的半定项$6x_1^2x_2^2$ 作放大估计. 简单把它们忽略掉就好了.

为找出半定项, 我们构造如下矩阵

$$V_k=[V_k^1,V_k^2,\cdots,V_k^n].$$

那么, V_k 的(i,j) 元即x^j 在x^k 的第i 个分量的指数. 例如, 在例8.4.6 中, 我们有

$$V_4=\begin{bmatrix} 4&3&3&2&3&2&2&1&3&2&2&1&2&1&1&0\\ 0&1&1&2&1&2&2&3&1&2&2&3&2&3&3&4 \end{bmatrix}^{\mathrm{T}}.$$

那么显然, 如果V_k 在第i 行的所有元素均为偶数, 则$P(x)=Fx^k$ 的第i 项所有变量均有偶指数, 我们称这种项为偶次方项. 而这时如果F 的相应元素$F_i<0$ 则该项负定. 那么, 在不等式估计中, 该项即可略去. 这样, 我们可以对$\tilde{F}$ 的定义作如下修正:

$$\tilde{F}_i=\begin{cases} F_i, & i\in D_n^k;\\ 0, & i\notin D_n^k,\ \text{偶次方项},\ F_i<0;\\ |F_i|, & \text{其他}. \end{cases} \tag{8.4.6}$$

今后, 我们只以(8.4.6) 作为$\tilde{F}$ 的定义.

回到方程(8.3.8), 将其记作

$$\dot{x}=\begin{bmatrix} C^1x^{k_1}\\ C^2x^{k_2}\\ \vdots\\ C^nx^{k_n} \end{bmatrix}. \tag{8.4.7}$$

假定所有的k_i, $i=1,\cdots,n$ 均为奇数. 找一个整数$m>0$ 使得$2m>\max\{k_1,\cdots,k_n\}$. 因为$2m-k_i+1>0$ 为偶数, 记$2n_i=2m-k_i+1$, 这里$n_i>0$.

下面我们定义一个正定的Lyapunov 函数:

$$L(x) = a_1 x_1^{2n_1} + \cdots + a_n x_n^{2n_n}, \tag{8.4.8}$$

这里$a_i > 0, \forall i$. 考虑

$$q_i := \frac{\partial L}{\partial x_i} g_i.$$

我们要检验q_i 的负定性. 因为我们需要用g_i 的信息, 上述结果需要作一定的修正. 首先要注意的是x_i 的指数增加了$2m - k_i$, 我们将g_i 的指数向量$V_{k_i}^j$ 换为

$$W_i^j = V_{k_i}^j + (2m - k_i)\delta_n^j \mathbf{1}_{n^{k_i - 1}}, \quad j = 1, \cdots, n; \ i = 1, \cdots, n.$$

再定义指数阵如下

$$W_i := [W_i^1, W_i^2, \cdots, W_i^n].$$

q_i 中的s 项称为偶次方项, 如果W_i 的第s 行的元素均为偶数.

注意, 在q_i 中我们只有一个对角元, 即x_i^{2m}, 不可能有x_j^{2m}, $j \neq i$. 这个元出现在

$$d_i = (i-1)\frac{n^{k_i} - 1}{n - 1} + 1.$$

现在我们定义

$$\tilde{C}_j^i = \begin{cases} C_j^i, & j = d_i; \\ 0, & j \neq d_i, \text{偶次方项 (相对于} W_i), \ C_j^i < 0; \\ |C_j^i|, & \text{其他}, \quad i = 1, \cdots, n. \end{cases} \tag{8.4.9}$$

回到系统(8.2.1), 利用前面的记号和讨论, 我们有下面的结论.

定理 8.4.7　*系统(8.2.1) 在原点是渐近稳定的, 如果存在$a_i > 0$, $i = 1, \cdots, n$, 整数m 满足$2m > \max\limits_{i=1,\cdots,n}\{k_i\}$ 使得*

$$\sum_{i=1}^n \frac{1}{2m} a_i \tilde{C}^i W_{k_i}^j < 0, \quad j = 1, \cdots, n. \tag{8.4.10}$$

证明　定义一个导数齐次Lyapunov 函数$L(x)$, 如(8.4.8). 根据上面的讨论可知

$$\dot{L}(x)\Big|_{(8.4.7)} \leq \sum_{j=1}^n \sum_{i=1}^n a_i \tilde{C}^i W_{k_i}^j x_j^{2m} < 0. \tag{8.4.11}$$

由于(8.4.7) 是(8.2.1) 的近似系统, 根据命题8.3.5 即得结论. □

注 在文献[19] 中上述结果称为CRDDP (跨行对角占优原则). 粗略地说, 它要求对角元对所有行相应变量下占优. 下面这个简化结果称为DDP (对角占优原则).

推论 8.4.8 系统(8.2.1) 是渐近稳定的, 如果存在整数m 满足$2m > \max\limits_{i=1,\cdots,n}\{k_i\}$ 使得

$$\tilde{C}^j\delta_n^j\mathbf{1}_{n^{k_j-1}} < 0, \quad j=1,\cdots,n. \tag{8.4.12}$$

证明 在(8.4.11) 中令$m\to\infty$ 可得

$$\lim_{m\to\infty}\sum_{i=1}^{n}\frac{1}{2m}[\tilde{C}^iV_{k_i}^j+(2m-k_i)\delta_n^j\mathbf{1}_{n^{k_i-1}}]=a_j\tilde{C}^j\delta_n^j\mathbf{1}_{n^{k_i-1}}.$$

与(8.4.12) 相比可知, 如果(8.4.12) 成立, 则可选足够大的m, 使得(8.4.11) 中的系数均为负数. □

注 在应用推论8.4.8 时, 有一个问题, 即什么是偶次方项, 因为这里无法构造W_i. 其实, 判定是否偶次方项不必依靠W_i, 可直接利用V_i. q_i 中的第j 项是$\dot{L}(x)$ 中的偶次方项当且仅当, 在V_i 的第j 行中除第j 个元素为奇数外, 其余元素均为偶数.

下面给出一个应用的简单例子. 通过这个简单例子来解释前面的定义、概念和结论.

例 8.4.9 找出λ 的一个范围, 使得下述系统在原点的某个邻域渐近稳定:

$$\begin{cases}\dot{x}_1=\sin(x_1)-x_1\cos(2\lambda x_2),\\ \dot{x}_2=x_2^2\ln(1-x_2-x_3)+0.5x_2^2x_3,\\ \dot{x}_3=2x_3^3(1-\cosh(x_3-x_2))-1.1x_3^5.\end{cases} \tag{8.4.13}$$

对(8.4.13) 作Taylor 展开, 可得到它的近似系统为

$$\begin{cases}\dot{x}_1=-\frac{1}{6}x_1^3+2\lambda^2x_1x_2^2\ (:=g_1),\\ \dot{x}_2=-x_2^3-\frac{1}{2}x_2^2x_3\ (:=g_2),\\ \dot{x}_3=-2.1x_3^5+2x_3^4x_2-x_3^3x_2^2\ (:=g_3).\end{cases} \tag{8.4.14}$$

应用(8.4.12), 对$j=1$ 得不等式

$$-\frac{1}{6}+2\lambda^2<0;$$

对$j=2$ 得不等式

$$-1+\frac{1}{2}<0;$$

对$j=3$ 注意到根据前面的注可知$x_3^3x_2^2$ 对应导数中的偶次方项, 且其系数为负, 故可略去, 因此, 得不等式

$$-2.1+2<0.$$

综上, 根据DDP (对角占优原则), 当$-\frac{1}{6}+2\lambda^2<0$ 时, 即

$$|\lambda|<0.2886751346,$$

系统(8.4.14) 在原点近似稳定. 因此(8.4.13) 在原点渐近稳定.

下面利用CRDDP (跨行对角占优原则). 取$m=3$, $a_i=1$, $i=1,2,3$. 那么, (8.4.10) 变为

$$\begin{cases}\frac{1}{6}>2\lambda^2\left(\frac{4}{6}\right),\\ 1>\frac{2}{6}\left(2\lambda^2\right)+\frac{1}{2}\left(\frac{5}{6}\right)+2\left(\frac{1}{6}\right),\\ 2.1>2\left(\frac{5}{6}\right)+\frac{1}{2}\left(\frac{1}{6}\right).\end{cases}$$

其解为

$$|\lambda|<0.3535533906.$$

即此时系统(8.4.14) 在原点近似稳定, 或(8.4.13) 在原点渐近稳定. □

从上面的例子可以看出, DDP 方法比CRDDP 方法方便, 但CRDDP 方法比DDP 方法精确. 容易证明, 这个结论具有一般性.

本节给出的半张量积形式的公式, 用计算机编程实现极为方便.

8.5　零中心系统的镇定

本节给出设计中心流形镇定非线性系统的一个典型的设计方法.

考虑一个反馈等价意义下的标准形式仿射非线性系统

$$\begin{cases}\dot{x}=Ax+Bv; \quad x\in\mathbb{R}^s;\\ \dot{z}=q(z,x), \quad z\in\mathbb{R}^t.\end{cases} \tag{8.5.1}$$

不失一般性, 设线性部分(A,B) 具有Brunovskey 标准型, 即(8.5.1) 的线性部分可

表示为

$$\begin{cases}\dot{x}_1^i = x_2^i,\\ \vdots\\ \dot{x}_{s_i-1}^i = x_{s_i}^i,\\ \dot{x}_{s_i}^i = v_i, \quad x^i \in \mathbb{R}^{s_i}, \quad i = 1, \cdots, m.\end{cases} \tag{8.5.2}$$

受例8.3.1 的启发, 我们寻找下述形式的控制器镇定系统:

$$v_i = a_1^i x_1^i + \cdots + a_{s_i}^i x_{s_i}^i - a_1^i \sum_{j=2}^{h} P_j^i(z), \quad i = 1, \cdots, m. \tag{8.5.3}$$

这里$P_j^i(z)$ 为j 次齐次多项式, 常数a_j^i 这样选择, 使得反馈后的线性子系统的系统矩阵为Hurwitz 矩阵. 设中心流形由等式$x = h(z)$ 刻画. 我们用$x = \phi(z)$ 来逼近它, 这里

$$x^i = \phi^i(z) = \begin{bmatrix}\sum\limits_{j=2}^{h} P_j^i(z)\\ 0\\ \vdots\\ 0\end{bmatrix}, \quad i = 1, \cdots, m. \tag{8.5.4}$$

根据中心流形逼近定理可知, 如果

$$\frac{\partial \phi}{\partial z} q(z, \phi(z)) = O(\|z\|^{r+1}), \tag{8.5.5}$$

那么两函数之差(即逼近误差) 为

$$\|\phi(z) - h(z)\| = O(\|z\|^{r+1}),$$

这里$\phi(z) = (\phi^1(z), \cdots, \phi^m(z))$.

其次, 我们构造中心流形上的动态方程的近似方程, 得

$$\dot{z}_i = q_i(z, \phi(z)), \quad i = 1, \cdots, t. \tag{8.5.6}$$

记(8.5.6) 的近似方程为

$$\dot{z} = \begin{bmatrix} D^1 x^{k_1}\\ \vdots\\ D^n x^{k_n}\end{bmatrix}, \tag{8.5.7}$$

这里, $D^i \neq 0, i = 1, \cdots, n$. 如果$k_i, i = 1, \cdots, n$ 均为奇数, 那么, 我们可以得到以下的主要定理:

定理 8.5.1 考虑系统(8.5.1)~(8.5.2), 如果存在如前所述的$\phi(z)$, 使得
(1) (8.5.5) 成立;
(2)

$$q_i(z, \phi(z)) - q_i(z, \phi(z) + O(\|z\|^{r+1})) = O(\|z\|^{k_i+1}), \quad i = 1, \cdots, t. \tag{8.5.8}$$

(3) 存在一个导数齐次Lyapunov 函数$V(z) > 0$ 使得$\dot{V}|_{(8.5.7)} < 0$.
那么, 系统(8.5.1)~(8.5.2) 可被控制(8.5.3) 镇定.

注 (a) 实际上, 条件(3) 保证了(8.5.6) 的近似稳定性. 而在条件(1) 的假定下, 条件(2) 保证了中心流形上的真实动态系统的渐近稳定性.

(b) 当然, 条件(3) 可以改为一个更一般的说法, 即系统(8.5.7) 近似稳定.

下面给出一个简单例子描述这个设计过程.

例 8.5.2 考查下述系统

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = x_3, \\ \dot{x}_3 = u, \\ \dot{z}_1 = z_1^2 z_2 + z_1 x_1 + z_2 x_2, \\ \dot{z}_2 = z_2^3 x_1. \end{cases} \tag{8.5.9}$$

为镇定线性部分, 设取控制如下,

$$u = -x_1 - 3x_2 - 3x_3 + \phi(z).$$

这里未定部分$\phi(z)$ 为z 的非线性多项式, 不妨设

$$h(z) = \alpha z_1^2 + \beta z_1 z_2 + \gamma z_2^2.$$

根据(8.5.4), 我们用下述函数逼近中心流形

$$\begin{cases} x_1(z) = \alpha z_1^2 + \beta z_1 z_2 + \gamma z_2^2, \\ x_2(z) = 0, \\ x_3(z) = 0. \end{cases}$$

根据中心流形逼近定理, 逼近误差为

$$
\frac{\partial x}{\partial z}\begin{bmatrix} z_1^2z_2+z_1x_1(z)+z_2x_2(z) \\ z_2^3x_1(z) \end{bmatrix} - \begin{bmatrix} x_2(z) \\ x_3(z) \\ -x_1(z)-3(x_2(z)-3x_3(z)+h(z) \end{bmatrix}
$$

$$
= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 2\alpha z_1+\beta z_2 & \beta z_2+2\gamma z_3 \end{bmatrix}\begin{bmatrix} z_1^2z_2+z_1x_1(z)+z_2x_2(z) \\ z_2^3x_1(z) \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = O(\|z\|^4).
$$

现在考虑中心流形上的动态系统. 我们有

$$
\begin{cases} \dot{z}_1 = z_1^2z_2+z_1[x_1(z)+O(\|z\|^4)], \\ \dot{z}_2 = z_2^3[x_1(z)+O(\|z\|^4)]. \end{cases}
$$

即

$$
\begin{cases} \dot{z}_1 = z_1^2z_2+z_1[\alpha z_1^2+\beta z_1z_2+\gamma z_2^2]+O(\|z\|^5)], \\ \dot{z}_2 = z_2^3[\alpha z_1^2+\beta z_1z_2+\gamma z_2^2]+O(\|z\|^7)]. \end{cases}
$$

于是, 中心流形上的动态系统的近似系统为

$$
\begin{cases} \dot{z}_1 = z_1^2z_2+z_1[\alpha z_1^2+\beta z_1z_2+\gamma z_2^2], \\ \dot{z}_2 = z_2^3[\alpha z_1^2+\beta z_1z_2+\gamma z_2^2]. \end{cases} \tag{8.5.10}
$$

选择一个导数齐次Lyapunov 函数如下

$$
V = z_1^4+z_2^2.
$$

并设$\alpha=\gamma=-1$, $\beta=0$, 则可得到

$$
\dot{V}|_{(8.5.10)} \le -4z_1^6+4z_1^5z_2-2z_2^6 \le -4z_1^6+\frac{10}{3}z_1^6+\frac{2}{3}z_2^6-2z_2^6<0, \quad z\ne 0.
$$

因此可以知道, 控制

$$
u=-x_1-3x_2-3x_3-z_1^2-z_2^2
$$

使闭环系统在原点渐近稳定. □

8.6　注释与参考

用中心流形方法进行非线性系统镇定器设计最早出现于20 世纪80 年代, 如文献[2]. 开始时主要针对个别具体系统. 对于最小相位系统的统一的处理方法开始于文献[11] 等, 这部分结果亦可参见文献[68], [71]. 文献[19] 首次对非最小相位非线性控制系统提出一套系统的设计方法. 这种方法在文献[22] 中作了进一步的简化和推广. 本章主要结果来自文献[19] 和[22]. 但用半张量积简化了其中的一些条件与计算, 使结果便于用计算机进行设计和计算.

本章的结果是建立在Byrnes-Isidori 标准型的基础上的, 当系统不能化成这种标准型时, 文献[23] 提出一种更广泛的标准型, 称为泛 Byrnes-Isidori 标准型. 证明几乎所有的系统都能化成这种标准型. 同时, 文献[23] 证明了文献[19] 中提出的中心流形设计方法对这种一般形式仍然可行. 作为应用的例子, 平面中心的设计可见文献[45].

对不具仿射形式的一般非线性系统, 文献[30] 讨论了其基于中心流形的镇定器设计方法.

第 9 章　动态系统的对称性

动态系统在群作用下的对称性是物理和数学中的重要课题. 因为自然界中的许多系统都具有对称性, 另外考虑对称性还可以大大地简化系统研究. 对称性也是复杂性科学研究的一个重点课题. 许多复杂现象是由于对称性及对称原理而导致的. 当考虑动力系统的对称性时, 利用Taylor 展开, 群在多元多项式上的作用变为其基本框架, 因此, 利用半张量积来研究动态控制系统的对称性具有极大的优越性.

9.1　对称群的结构和它的李代数

本章考虑的对称性是关于$\mathbb{R}^n$ 上的动态系统在李群作用下的不变性. 设G 是一个李群, G 是$\mathbb{R}^n$ 上的一个群作用 (或者作用在$\mathbb{R}^n$ 的一个开子集上) , 如果存在一个映射θ: $G \times \mathbb{R}^n \to \mathbb{R}^n$ 使得(i) $\theta(e)\,x = x$, $\forall x \in \mathbb{R}^n$; (ii) 对于任意$\alpha_1, \alpha_2 \in G$, 有$\theta(\alpha_1\alpha_2)\,x = \theta(\alpha_1)\,(\theta(\alpha_2)\,x)$.

对于一个控制系统, 我们定义两种对称性如下:

定义 9.1.1　给定一个解析控制系统

$$\dot{x} = f_0(x) + \sum_{i=1}^{m} f_i(x)\,u_i, \qquad x \in \mathbb{R}^n, \tag{9.1.1}$$

其中$f_i(x)$, $i = 0, \cdots, m$ 是解析向量场. 设G 是作用在$\mathbb{R}^n$ 上的一个李群(或是作用在$\mathbb{R}^n$ 的一个开子集$M \subset \mathbb{R}^n$ 上) .

(1) 称系统(9.1.1) 关于G 状态空间对称 (或有一个状态空间对称群G) , 如果对每个$\alpha \in G$ 有

$$\theta(\alpha)_* f_i(x) = f_i(\theta(\alpha)\,x), \qquad i = 0, \cdots m,$$

其中$\theta(\alpha)_*$ 是$\theta(\alpha)$ 的诱导映射, 后者是$\mathbb{R}^n$ 上的一个微分同胚. 如果$f_i(x)$ 满足上面的方程(对一个给定的α) , 称$f_i(x)$ 是$\theta(\alpha)$ 不变的.

(2) 称系统(9.1.1) 关于G 对称, 简称系统对称 (或有一个对称群G) , 如果对每个$\alpha \in G$ 有

$$\theta(\alpha)_* \mathcal{A} = \mathcal{A},$$

其中

$$\mathcal{A} = \left\{ f(x) + \sum_{i=1}^{m} g_i(x) u_i \;\middle|\; u \in \mathbb{R}^m \right\}.$$

如果G 是一般线性群的子群, 即$G < GL(n,\mathbb{R})$, 那么称系统(9.1.1) 关于G 线性状态空间对称 (或有一个线性状态空间对称群G) .

我们从系统(9.1.1) 的线性状态空间对称开始. 对于状态空间对称, 控制不是本质的, 因此我们考虑自由解析系统

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n. \tag{9.1.2}$$

利用Taylor 级数展开, 我们可以将f 表示成

$$f(x) = f_0 + f_1 x + f_2 x^2 + \cdots, \tag{9.1.3}$$

其中$f_i \in M_{n\times n^i}$, $x = (x_1, \cdots, x_n)^{\mathrm{T}}$, 并且所有的乘法都是左半张量积.

设$\alpha \in GL(n,\mathbb{R})$, $\theta_\alpha : x \mapsto y = \alpha x$, 那么$f$ 在θ_α 下不变意味着

$$(\theta_\alpha)_*(f(x)) = \alpha f(x) = \alpha f(\alpha^{-1}y) = f(y), \quad \forall y \in \mathbb{R}^n. \tag{9.1.4}$$

它等价于

$$\alpha f(x) = f(\alpha x), \quad \forall x \in \mathbb{R}^n. \tag{9.1.5}$$

由于$(\theta_\alpha)_*$ 不改变一个齐次向量场的阶数, 如果(9.1.4) 对f 成立, 它应该对f 的每个齐次项也成立, 即

$$\alpha f_k x^k = f_k(\alpha x)^k, \quad \forall x \in \mathbb{R}^n;\ k = 0, 1, \cdots. \tag{9.1.6}$$

利用左半张量积的性质, 我们有

$$\begin{aligned}
(\alpha x)^k &= \underbrace{\alpha x \ltimes \alpha x \ltimes \cdots \ltimes \alpha x}_{k} \\
&= \alpha(I_n \otimes \alpha)x^2 \underbrace{\alpha x \ltimes \alpha x \ltimes \cdots \ltimes \alpha x}_{k-2} \\
&= \cdots \\
&= (\alpha)(I_n \otimes \alpha)(I_{n^2} \otimes \alpha)\cdots(I_{n^{k-1}} \otimes \alpha)x^k \\
&= (\alpha \otimes I_n)(I_n \otimes \alpha)[(I_{n^2} \otimes \alpha)\cdots(I_{n^{k-1}} \otimes \alpha)]x^k \\
&= (\alpha \otimes \alpha)[(I_{n^2} \otimes \alpha)\cdots(I_{n^{k-1}} \otimes \alpha)]x^k \\
&= \cdots \\
&= \underbrace{(\alpha \otimes \alpha \otimes \cdots \otimes \alpha)}_{k} x^k := \alpha^{\otimes k} x^k.
\end{aligned}$$

显然, 系统(9.1.2) 是α-不变的, 当且仅当

$$\alpha f_k x^k = f_k \alpha^{\otimes k} x^k, \quad k = 0, 1, \cdots. \tag{9.1.7}$$

回忆第 5 章, 由于x^k 是k 次齐次多项式的冗余基, 即它包括一组基底. 为了避免冗余, 我们使用一组标准基底, 记这组基底为$x_{(k)}$, 它是一个集合

$$x_{(k)} = \left\{ \prod_{i=1}^{n} x_i^{t_i} \,\middle|\, \sum_{i=1}^{n} t_i = k \right\}.$$

$x_{(k)}$ 也可看作一个矩阵. $x_{(k)}$ 里的元素按照字母序排列, 即, 设$b_1 = x_1^{p_1} \cdots x_n^{p_n}$, $b_2 = x_1^{q_1} \cdots x_n^{q_n}$, 定义$b_1 \prec b_2$, 如果$p_s = q_s, s = 1, \cdots, t$, 且对某个$0 \le t < n$, $p_{t+1} > q_{t+1}$. 因此, 当将$x_{(k)}$ 看作一个矩阵时, 它可以表示成$x_{(k)} = \begin{bmatrix} b_1, & \cdots, & b_d \end{bmatrix}^{\mathrm{T}}$.

容易验证, 对于$x \in \mathbb{R}^n$, k 次齐次多项式组成的向量空间的维数是

$$d := r_n^k = \frac{(n+k-1)!}{(n-1)!k!}. \tag{9.1.8}$$

我们用一个简单的例子来描述冗余基底x^k 和基底$x_{(k)}$.

例 9.1.2 设$n = 2$, $k = 3$, 那么$x = (x_1, x_2)^{\mathrm{T}}$, 且$d = (2+3-1)!/3! = 4$.

$$x^3 = \begin{bmatrix} x_1^3 & x_1^2 x_2 & x_1 x_2 x_1 & x_1 x_2^2 & x_2 x_1^2 & x_2 x_1 x_2 & x_2^2 x_1 & x_2^3 \end{bmatrix}^{\mathrm{T}},$$

且

$$x_{(3)} = \begin{bmatrix} x_1^3 & x_1^2 x_2 & x_1 x_2^2 & x_2^3 \end{bmatrix}^{\mathrm{T}}.$$

□

于是我们可以构造一个矩阵$T_N(n, k) \in M_{n^k \times r_n^k}$ 使得

$$x^k = T_N(n, k) x_{(k)}. \tag{9.1.9}$$

实际上, 我们在第5 章曾经给出$T_N(n, k)$ 的具体构造.

由于一个基底的系数是唯一的, 由(9.1.7) 我们有

命题 9.1.3 系统(9.1.2) 是α-不变的, 当且仅当

$$\alpha f_k T_N(n, k) = f_k \alpha^{\otimes k} T_N(n, k), \quad k = 0, 1, 2, \cdots, \tag{9.1.10}$$

这里

$$\alpha^{\otimes k} = \underbrace{\alpha \otimes \alpha \otimes \cdots \otimes \alpha}_{k}.$$

显然, f 是α-不变的一个充分条件是

$$\alpha f_k = f_k \alpha^{\otimes k}, \quad k = 0, 1, 2, \cdots. \tag{9.1.11}$$

利用命题9.1.3, 我们可以直接得到下面的结论.

命题 9.1.4　*设H 是$GL(n,\mathbb{R})$ 的一个子集, 其中$GL(n,\mathbb{R})$ 由所有满足(9.1.7) 的α 组成, 则H 是一个群.*

方程(9.1.10) 给出了α 满足的代数等式, 但是一般很难解出这样的无穷个代数方程. 我们需要找到一个简单的方法来解决这个问题, 于是我们转向李代数方法.

记$g(G)$ 是G 的李代数, G 是$gl(n,\mathbb{R})$ 的一个李子代数. 我们证明下面这个基本的引理.

引理 9.1.5　*设$G < GL(n,\mathbb{R})$ 是一个连通李群. 系统(9.1.2) (或者简单地说, 向量场$f(x)$) 具有对称群G, 当且仅当*

$$\mathrm{ad}_{Vx} f(x) = 0, \quad \forall V \in g(G), \tag{9.1.12}$$

其中Vx 是一个线性向量场.

证明　设M 是一个给定的流形. 对于一个向量场$X \in V(M)$, 我们记它的以$x(0)=x$ 为初值的积分曲线为$\phi_t^X(x)$, 于是对于任意$Y \in V(M)$ 有

$$(\phi_t^X)_* Y(x) = Y(\phi_t^X(x)).$$

当且仅当$[X,Y]=0$ [10]. 这样$Vx \in V(\mathbb{R}^n)$ 的积分曲线是$e^{Vt}x$, 因此

$$(e^{Vt})_* f(x) = e^{Vt} f(x(z)) = e^{Vt} f(e^{-Vt}z) = f(z),$$

其中$z = e^{Vt}x$. 等价地

$$e^{Vt} f(x) = f(e^{Vt}x),$$

当且仅当$\mathrm{ad}_{Vx} f(x) = 0$. □

注意在引理9.1.5 和本章后面的讨论中不包括离散群.

考虑系统(9.1.1). 对每个f_j, $j = 0, 1, \cdots, m$ 利用Taylor 级数展开, 我们记

$$f_j = \sum_{k=0}^{\infty} f_k^j x^k, \quad i = 0, \cdots, m.$$

由于$\deg(ad_{Vx}f_k^jx^k)=k$, 即ad_{Vx} 不改变每项的次数, 我们可以定义

$$\mathcal{V}_k^j=\{V\in gl(n,\mathbb{R})\mid \mathrm{ad}_{Vx}f_k^jx^k=0\}.$$

利用Jacobi 恒等式, 容易看出$\mathcal{V}_k^j$ 是一个李代数. 根据引理9.1.5, 如果$G<GL(n,\mathbb{R})$ 是系统(9.1.1) 的最大状态空间对称群, 那么它的李代数是

$$g(G):=\mathcal{V}=\bigcap_{j=0}^{m}\bigcap_{k=0}^{\infty}\mathcal{V}_k^j.$$

于是可以构造以g 为李代数的相应的连通群$G(g)$ 为

$$G(g)=\left\{\prod_{i=1}^{k}exp(t_iV_i)\,\middle|\, V_i\in\mathcal{V},\ k<\infty\right\}. \tag{9.1.13}$$

综上所述, 得到下面的结论.

定理 9.1.6 系统(9.1.1) 具有唯一最大连通状态空间对称群$G<GL(n,\mathbb{R})$, 且它的李代数是

$$g(G)=\bigcap_{j=0}^{m}\bigcap_{k=0}^{\infty}\mathcal{V}_k^j. \tag{9.1.14}$$

最后, 假设给定一个李代数$g\subset gl(n,\mathbb{R})$, 对于以$G(g)$ 作为状态空间对称群的向量场$f(x)$ 的集合, 我们给出一个代数条件.

记$\mathcal{H}_n^k$ 为分量是k 次齐次多项式的向量场全体的集合. 容易看出$\mathcal{H}_n^k$ 是$\mathbb{R}$ 上的线性空间, 并且对任意$V\in gl(n,\mathbb{R})$, 映射$\mathrm{ad}_{Vx}:\mathcal{H}_n^k\to\mathcal{H}_n^k$ 都是线性映射. 利用(9.1.8), $\mathcal{H}_n^k$ 的维数$d_n^k=n(n+k-1)!/(n-1)!k!$. 于是可以得到$\mathcal{H}_n^k$ 的一组基底, 记为H_n^k 的列, 为

$$H_n^k=I_n\otimes x_{(k)}^{\mathrm{T}}, \tag{9.1.15}$$

称之为$\mathcal{H}_n^k$ 的自然基底. 在后面ad_{Vx} 的伴随表现意味着关于这组自然基底的表现.

现在我们可以定义一个从$gl(n,\mathbb{R})$ 到李导数的伴随表现的映射, 称之为伴随映射.

定义 9.1.7 定义伴随映射如下:

$$\Phi_n^k:gl(n,\mathbb{R})\ni V\to\Phi_n^k(V)\in gl(d_n^k,\mathbb{R}).$$

其中$\Phi_n^k(V)$ 是$\mathrm{ad}_{Vx}:\mathcal{H}_n^k\to\mathcal{H}_n^k$ (关于自然基底) 的伴随表现.

我们给出一个例子来说明.

例 9.1.8　设$n=2$, 且

$$V=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}. \tag{9.1.16}$$

直接计算有

$$\Phi_2^k(V)=\begin{bmatrix}A & -I\\ 0 & A\end{bmatrix}. \tag{9.1.17}$$

其中

$$A=\begin{bmatrix}0 & & & & \\ k & 0 & & & \\ & k-1 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 1 & 0\end{bmatrix}. \tag{9.1.18}$$

□

下面的命题是引理9.1.5 和Φ_n^k 的定义的一个直接结果.

命题 9.1.9　设$G\in GL(n,\mathbb{R})$ 是一维连通子群, 且$V\in g(G)$. 一个分量为k 次齐次多项式的向量场$f(x)$ 是G-不变的, 当且仅当, $f(x)\in\ker(\Phi_n^k(V))$.

例 9.1.10　回顾例9.1.8.　考虑$\ker(\Phi_n^k(V))$, 其中V 由(9.1.16) 给定.　利用(9.1.17), 我们要解下面的方程

$$\begin{bmatrix}A & -I\\ 0 & A\end{bmatrix}\begin{bmatrix}X\\ Y\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}.$$

于是我们有

$$\begin{cases}Y=AX,\\ A^2X=0.\end{cases}$$

利用(9.1.18) 有

$$A^2 = \begin{bmatrix} 0 & & & & \\ 0 & 0 & & & \\ k(k-1) & 0 & & & \\ & \ddots & \ddots & & \\ 0 & & 2\times 1 & 0 & 0 \end{bmatrix}.$$

因此$X = (0, \cdots, a, b)^{\mathrm{T}}$, $Y = AX = (0, \cdots, 0, a)^{\mathrm{T}}$, 其中$a$ 和b 是任意两个实数. 由于$(col(X), col(Y))^{\mathrm{T}}$ 是关于$\mathcal{H}_2^k$ 的自然基底的系数, 于是有

$$f(x) = (ax_1x_2^{k-1} + bx_2^k, ax_2^k)^{\mathrm{T}}, \quad k \geq 1. \tag{9.1.19}$$

根据命题9.1.9, 我们得出结论, 这样的向量场$f(x)$ 具有一个一维对称群

$$G = exp\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} t\right). \tag{9.1.20}$$

□

9.2 旋转下的对称性

本节考虑旋转变换下的状态空间对称. 旋转对称由于其明确的物理意义, 在对称性研究中具有特殊的重要性.

考虑系统(9.1.1), 并假设$n = 2$, 于是下面的定理回答了何时它具有状态空间对称群$SO(2, \mathbb{R})$.

定理 9.2.1 [114] 当$n = 2$ 时系统(9.1.1) 具有状态空间对称群$SO(2, \mathbb{R})$, 当且仅当

$$f_j(x) = \sum_{i=0}^{\infty} (x_1^2 + x_2^2)^i \begin{bmatrix} a_i^j & b_i^j \\ -b_i^j & a_i^j \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad a_i^j, b_i^j \in \mathbb{R},\ j = 0, \cdots, m. \tag{9.2.1}$$

我们考虑何时系统(9.1.1) 具有一个状态空间对称群$SO(n, \mathbb{R})$, 这个问题是文献[114] 的一个推广. 我们的充要条件如下:

定理 9.2.2 系统(9.1.1) 在$n \geq 3$ 时具有一个状态空间对称群$G = SO(n, \mathbb{R})$, 当且仅当

$$f_j(x) = \sum_{i=0}^{\infty} a_i^j ||x||^{2i} x, \quad a_i^j \in \mathbb{R}, \quad j = 0, 1, \cdots, m, \tag{9.2.2}$$

这里$||x||^{2i} = (x_1^2 + \cdots + x_n^2)^i$.

为了证明定理9.2.2, 我们需要做一些准备.

定义 9.2.3　称一个光滑函数$p(x) \in V(\mathbb{R}^n)$ 是$A \in M_{n\times n}$ 不变的, 如果

$$L_{Ax}p(x) = 0. \tag{9.2.3}$$

"不变"的意思来自于下面的事实: 由于$L_{Ax}p(x) \equiv 0$, 则$L^k_{Ax}p(x) = 0$, $k > 1$. 利用Taylor 级数展开, 我们有

$$(\phi_t^{Ax})^* p(x) = p(e^{At}x) = \sum_{k=0}^{\infty} L^k_{Ax}p(x)\frac{t^k}{k!} = p(x).$$

即$p(x)$ 关于Ax 的积分曲线是不变的.

在一些文献中, $p(x)$ 也被称为线性向量场Ax 的一次积分.

引理 9.2.4　设$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, 则$A$ 无奇数阶的不变多项式.

证明　我们只需证明关于齐次多项式部分. 设$g(y_1, y_2) = \sum_{i=0}^{2l+1} a_i y_1^i y_2^{2l+1-i}$, 且$L_{Ay}g(y_1, y_2) = 0$, 那么

$$\begin{aligned} 0 &= L_{Ay}g(y_1, y_2) \\ &= \frac{\partial g}{\partial y}\begin{bmatrix} -y_2 \\ y_1 \end{bmatrix} = -\sum_{i=0}^{2l+1} i a_i y_1^{i-1} y_2^{2l-i+2} + \sum_{j=0}^{2l+1}(2l+1-j)a_j y_1^{j+1} y_2^{2l-j} \\ &= -\sum_{j=0}^{2l-1}(j+2)a_{j+2}y_1^{j+1}y_2^{2l-j} \\ &\quad + \sum_{j=0}^{2l-1}(2l+1-j)a_j y_1^{j+1} y_2^{2l-j} - a_1 y_2^{2l+1} + a_{2l} y_1^{2l+1}. \end{aligned}$$

比较两边的系数得到

$$a_1 = 0, \quad a_{2l} = 0, \quad (j+2)a_{j+2} = (2l+1-j)a_j, \quad j = 1, 2, \cdots, 2l-1.$$

因此

$$a_i = 0, \quad 0 \le i \le 2l+1.$$

□

引理 9.2.5　设A 如引理9.2.4 所示, 则A 没有关于两个变量均为奇数阶的不变多项式.

证明 令$g(y_1,y_2)=\sum\limits_{i=0}^{m-1}a_{2i+1}y_1^{2i+1}y_2^{2m-2i-1}$, 则

$$\begin{aligned}
0 &= L_{Ay}g(y_1,y_2)=\tfrac{\partial g}{\partial y}\begin{bmatrix}-y_2\\ y_1\end{bmatrix}\\
&= -\sum_{i=0}^{m-1}(2i+1)a_{2i+1}y_1^{2i}y_2^{2m-2i}+\sum_{j=0}^{m-1}(2m-2j-1)a_{2j+1}y_1^{2j+2}y_2^{2m-2j-2}\\
&= -\sum_{i=1}^{m-1}(2i+1)a_{2i+1}y_1^{2i}y_2^{2m-2i}\\
&\quad+\sum_{i=1}^{m-1}(2m-2i+1)a_{2i-1}y_1^{2i}y_2^{2m-2i}-a_1y_2^{2m}+a_{2m-1}y_1^{2m}.
\end{aligned}$$

比较两边系数得到

$$a_1=0,\quad a_{2m-1}=0,\quad (2i+1)a_{2i+1}=(2m-2i+1)a_{2i-1},\quad i=1,2,\cdots,m-1,$$

即有

$$a_{2i+1}=0,\quad 0\le i\le m-1.$$

□

引理 9.2.6 考虑三阶正交代数$o(3,\mathbb{R})$ 和多项式

$$g(x_1,x_2,x_3)=\sum_{i_1+i_2+i_3=2k}a_{i_1i_2i_3}x_1^{i_1}x_2^{i_2}x_3^{i_3}.$$

如果$g(x)$ 是$o(3,\mathbb{R})$ 不变的, 那么对于含有至少一个i_1, i_2 或i_3 为奇数的项, 我们有

$$a_{i_1i_2i_3}=0.$$

证明 设$\{v_1,v_2,v_3\}$ 是$o(3,\mathbb{R})$ 的一组标准基, 这里

$$v_1=\begin{bmatrix}0&0&0\\0&0&-1\\0&1&0\end{bmatrix},v_2=\begin{bmatrix}0&0&1\\0&0&0\\-1&0&0\end{bmatrix},v_3=\begin{bmatrix}0&-1&0\\1&0&0\\0&0&0\end{bmatrix}.$$

于是$L_{v_ix}g=0$, $i=1,2,3$.

假设i_1 是奇数, 由$L_{v_1x}g=0$, 我们有

$$\frac{\partial g}{\partial x}\begin{bmatrix}0\\-x_3\\x_2\end{bmatrix}=0.$$

利用引理9.2.4 我们有$a_{i_1i_2i_3}=0$.

同理, 当i_2 或i_3 是奇数时, 我们也有$a_{i_1i_2i_3}=0$. □

引理 9.2.7　设$f\in\mathcal{H}_3^{2k}$ 表示为

$$f(x)=\begin{bmatrix}f_1\\f_2\\f_3\end{bmatrix},\quad f_i=\sum_{i_1+i_2+i_3=2k}a^i_{i_1i_2i_3}x_1^{i_1}x_2^{i_2}x_3^{i_3},\quad i=1,2,3.$$

假设它是$o(3,\mathbb{R})$ 不变的, 即

$$[v_ix,f]=\frac{\partial f}{\partial x}v_ix-v_if=0,\quad i=1,2,3.$$

则$f(x)\equiv 0$.

证明　由于

$$\frac{\partial f}{\partial x}v_ix=v_if,\quad i=1,2,3,$$

直接计算可得到下面的(9.2.4) 和(9.2.5).

$$\frac{\partial f_1}{\partial x}\begin{bmatrix}0\\-x_3\\x_2\end{bmatrix}=\frac{\partial f_2}{\partial x}\begin{bmatrix}x_3\\0\\-x_1\end{bmatrix}=\frac{\partial f_3}{\partial x}\begin{bmatrix}-x_2\\x_1\\0\end{bmatrix}=0,\tag{9.2.4}$$

$$\frac{\partial f_2}{\partial x}\begin{bmatrix}0\\-x_3\\x_2\end{bmatrix}=-f_3,\quad \frac{\partial f_3}{\partial x}\begin{bmatrix}0\\-x_3\\x_2\end{bmatrix}=f_2,\quad \frac{\partial f_3}{\partial x}\begin{bmatrix}x_3\\0\\-x_1\end{bmatrix}=-f_1.\tag{9.2.5}$$

考虑(9.2.4). 根据引理9.2.6, f_i 中非零项的每个变量都是偶数次的.

观察(9.2.5), 在它的左边每项至少有一个变量是奇数次的, 而在它的右边所有的变量都是偶数次的. 于是有

$$f_1=f_2=f_3\equiv 0,\quad \Rightarrow f\equiv 0.$$

□

引理 9.2.8　给定一个多项式

$$g(x_1,x_2,x_3)=\sum_{i_1+i_2+i_3=2k+1}a_{i_1i_2i_3}x_1^{i_1}x_2^{i_2}x_3^{i_3},$$

并且假设$L_{v_1 x}g=0$, 则

$$g=\left(\sum_{j_1+j_2+j_3=k} b_{j_1j_2j_3}x_1^{2j_1}x_2^{2j_2}x_3^{2j_3}\right)x_1. \tag{9.2.6}$$

证明 由于$L_{v_1 x}g=0$, 我们有

$$\frac{\partial g}{\partial x}\begin{bmatrix}0\\-x_3\\x_2\end{bmatrix}=0.$$

如果i_1 是偶数, 那么i_2+i_3 是奇数. 由引理9.2.6 我们有$a_{i_1i_2i_3}=0$.

如果i_1 是奇数, 那么i_2 和i_3 要么都是奇数, 要么都是偶数. 对于第一种情况, 根据引理9.2.5, $a_{i_1i_2i_3}=0$;

对于第二种情况, 假设

$$b_{j_1j_2j_3}=a_{(2j_1+1)(2j_2)(2j_3)},\quad j_1+j_2+j_3=k.$$

于是结论成立. □

注 如果$L_{v_i x}g=0$, $i=2$ 或$i=3$, 由类似的讨论有

$$g=\left(\sum_{j_1+j_2+j_3=k} b_{j_1j_2j_3}x_1^{2j_1}x_2^{2j_2}x_3^{2j_3}\right)x_i.$$

引理 9.2.9 设$f(x)\in\mathcal{H}_3^{2k+1}$ 表示为

$$f(x)=\begin{bmatrix}f_1\\f_2\\f_3\end{bmatrix},$$

其中f_1,f_2,f_3 是$2k+1$ 次齐次多项式. 如果

$$\frac{\partial f}{\partial x}v_i x=v_i f,\quad i=1,2,3,$$

那么

$$
\left\{
\begin{aligned}
f_1 &= \left(\sum_{j_1+j_2+j_3=k} a_{j_1j_2j_3} x_1^{2j_1} x_2^{2j_2} x_3^{2j_3} \right) x_1, \\
f_2 &= \left(\sum_{j_1+j_2+j_3=k} b_{j_1j_2j_3} x_1^{2j_1} x_2^{2j_2} x_3^{2j_3} \right) x_2, \\
f_3 &= \left(\sum_{j_1+j_2+j_3=k} c_{j_1j_2j_3} x_1^{2j_1} x_2^{2j_2} x_3^{2j_3} \right) x_3,
\end{aligned}
\right.
\tag{9.2.7}
$$

而且, 如果$a_{k00} = 0$, 那么

$$
a_{j_1j_2j_3} = 0, \quad b_{j_1j_2j_3} = 0, \quad c_{j_1j_2j_3} = 0.
$$

证明 由方程(9.2.4) 和引理9.2.8 我们有(9.2.7), 再利用

$$
\frac{\partial f(x)}{\partial x} v_3 x = v_3 f(x),
$$

可以直接推出下面的(9.2.8) 和(9.2.9).

$$
\frac{\partial f_2}{\partial x} \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix} = f_1, \tag{9.2.8}
$$

$$
\frac{\partial f_1}{\partial x} \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix} = -f_2. \tag{9.2.9}
$$

由(9.2.8) 和(9.2.9) 分别有

$$
\begin{aligned}
&\sum_{j_1+j_2+j_3=k} -b_{j_1j_2j_3}(2j_1) x_1^{2j_1-1} x_2^{2j_2+2} x_3^{2j_3} \\
&+ \sum_{j_1+j_2+j_3=k} b_{j_1j_2j_3}(2j_2+1) x_1^{2j_1+1} x_2^{2j_2} x_3^{2j_3} \\
= &\sum_{i_1+i_2+i_3=k} a_{i_1i_2i_3} x_1^{2i_1+1} x_2^{2i_2} x_3^{2i_3} \\
= &\sum_{\substack{i_1+i_2+i_3=k \\ i_1 \le k-1,\ i_2 \ge 1}} [-2(i_1+1) b_{(i_1+1)(i_2-1)i_3} + (2i_2+1) b_{i_1i_2i_3}] x_1^{2i_1+1} x_2^{2i_2} x_3^{2i_3} \\
&+ b_{k00} x_1^{2k+1};
\end{aligned}
\tag{9.2.10}
$$

和

$$
\begin{aligned}
& - \sum_{j_1+j_2+j_3=k} (2j_1+1)a_{j_1j_2j_3}x_1^{2j_1}x_2^{2j_2+1}x_3^{2j_3} \\
& + \sum_{j_1+j_2+j_3=k} (2j_2)a_{j_1j_2j_3}x_1^{2j_1+2}x_2^{2j_2-1}x_3^{2j_3} \\
= & - \sum_{i_1+i_2+i_3=k} b_{i_1i_2i_3}x_1^{2i_1}x_2^{2i_2+1}x_3^{2i_3} \\
= & \sum_{\substack{i_1+i_2+i_3=k \\ i_1\geq 1,\ i_2\leq k-1}} [-(2i_1+1)a_{i_1i_2i_3}+2(i_2+1)a_{(i_1-1)(i_2+1)i_3}]x_1^{2i_1}x_2^{2i_2+1}x_3^{2i_3} \\
& + a_{0k0}x_2^{2k+1}.
\end{aligned} \tag{9.2.11}
$$

比较(9.2.10) 两边的系数, 我们有

$$
\begin{cases}
(2i_2+1)b_{i_1i_2i_3}-2(i_1+1)b_{(i_1+1)(i_2-1)i_3}=a_{i_1i_2i_3}, \\
i_1+i_2+i_3=k, i_1\leq k-1, i_2\geq 1, \\
b_{k00}=a_{k00}=0.
\end{cases} \tag{9.2.12}
$$

同理, 由(9.2.11) 有

$$
\begin{cases}
(2i_1+1)a_{i_1i_2i_3}-2(i_2+1)a_{(i_1-1)(i_2+1)i_3}=b_{i_1i_2i_3}, \\
i_1+i_2+i_3=k, i_1\geq 1, i_2\leq k-1, \\
b_{0k0}=a_{0k0}.
\end{cases} \tag{9.2.13}
$$

因此

$$
\begin{cases}
2(i_2+1)a_{(i_1-1)(i_2+1)i_3}=(2i_1+1)a_{i_1i_2i_3}-b_{i_1i_2i_3}, \\
(2i_2+1)b_{i_1i_2i_3}=a_{i_1i_2i_3}+2(i_1+1)b_{(i_1+1)(i_2-1)i_3}, \\
a_{k00}=b_{k00}=0.
\end{cases} \tag{9.2.14}
$$

于是有$a_{i_1i_20}=b_{i_1i_20}=0,\quad i_1+i_2=k.$

同理, 我们有

$$2(i_3+1)a_{(i_1-1)i_2(i_3+1)}=(2i_1+1)a_{i_1i_2i_3}-c_{i_1i_2i_3},$$

和

$$(2i_3+1)c_{i_1i_2i_3}=a_{i_1i_2i_3}+(2i_1+1)c_{(i_1+1)i_2(i_3-1)}.$$

因此

$$a_{i_1i_2i_3} = c_{i_1i_2i_3} = 0, \quad i_1 + i_3 = k - i_2 \text{ 或 } i_1 + i_2 + i_3 = k.$$

同理, 我们也有

$$b_{i_1i_2i_3} = 0.$$

□

下面的引理是文献[114] 中主要结果的一个一般化.

引理 9.2.10 考虑系统

$$\dot{x} = f(x) = \sum_{i=1}^{t} p_i(x)K_i x, \quad x \in \mathbb{R}^n,\ t \in \mathbb{Z}_+, \tag{9.2.15}$$

其中$p_i(x)$ 是一个多项式, 且$K_i \in M_{n\times n}$. 系统(9.2.15) 关于$G < GL(n,\mathbb{R})$ 状态空间对称, 如果

(1) $p_i(x)$, $i = 1,\cdots,t$ 是$g(G)$ 不变的;

(2) K_i, $i = 1,\cdots,t$ 是$g(G)$ 的中心[99], 其中$g(G)$ 是G 的李代数.

证明 设$V \in g(G)$, 则

$$\begin{aligned} \mathrm{ad}_{Vx} f(x) &= \sum_{i=1}^{t}(L_{Vx}p_i(x)K_i x + p_i(x)ad_{Vx}K_i x) \\ &= \sum_{i=1}^{t}(L_{Vx}p_i(x)K_i x - p_i(x)[V, K_i]x) = 0. \end{aligned}$$

由引理9.1.5 知结论成立. □

引理 9.2.11 当$n = 3$ 时系统(9.1.1) 有一个状态空间对称群$G = SO(3,\mathbb{R})$, 当且仅当

$$f_j(x) = \sum_{i=0}^{\infty} a_k^j \|x\|^{2i} x, \quad a_k^j \in \mathbb{R},\ j = 0, 1, \cdots, m. \tag{9.2.16}$$

证明 充分性. 由引理9.2.10 直接得到.

必要性. 考虑$n = 3$ 时的系统(9.1.1), 假设它是关于$G = SO(3,\mathbb{R})$ 状态空间对称, 并且

$$f_j(x) = \sum_{r=0}^{\infty} f_r^j(x),$$

其中$f_r^j(x) \in \mathcal{H}_n^r$. 如果$r$ 是偶数, 根据引理9.2.7 有$f_r^j = 0$. 因此我们假设$r = 2k+1$. 记$f_{2k+1}^j(x)$ 中x_1^{2k+1} 的系数为$a_k = a_{k00}$, 令

$$g_{2k+1}^j(x) = f_{2k+1}^j(x) - a_k||x||^{2k}x.$$

根据引理9.1.5, 直接计算有

$$\frac{\partial g_{2k+1}}{\partial x}v_i x = v_i g_{2k+1}, \quad i = 1, 2, 3.$$

引理9.2.8 确保

$$g_{2k+1}(x) \equiv 0.$$

于是有

$$f_{2k+1}(x) = a_k||x||^{2k}x.$$

□

定理9.2.2 的证明 从引理9.2.11 的证明中可以看出, 证明中使用的基本技巧是比较一对变量. 显然这个方法可以推广到$n > 3$ 的情形, 于是可以得到定理9.2.2. □

注 比较定理9.2.2 和定理9.2.1, 可以看出对于$n = 2$ 和$n \geq 3$, 相应的$f(x)$ 完全不同. 一个原因可能是由于它们李代数的结构, $o(2, \mathbb{R})$ 的中心是

$$Z(o(2, \mathbb{R})) = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \middle| a, b \in \mathbb{R} \right\},$$

而$o(n, \mathbb{R})$, $n \geq 3$ 的中心是

$$Z(o(n, \mathbb{R})) = \{rI_n \mid r \in \mathbb{R}\}.$$

它们完全不同. 当$n \geq 3$ 时$o(n, \mathbb{R})$ 没有非平凡中心. 粗略地说, 就是没有"交换"的自由. 为了读者的方便, 我们回忆一下, 一个李代数L 的中心是[99]

$$Z = \{z \in L \parallel [z, l] = 0, \ \forall l \in L\}.$$

9.3 平面系统的对称性

本节考虑平面系统的状态空间对称. 下面的主要结果刻画了所有可能的状态空间对称.

定理 9.3.1　(1) 假设$n=2$时系统(9.1.1)有一个连通的状态空间对称群$G<GL(2,\mathbb{R})$，那么G一定与下面四个群中的一个共轭.

$$G_1=\left\{exp\left(\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}t\right)\,\middle|\, t\in\mathbb{R}\right\}, \tag{9.3.1}$$

且

$$\frac{\lambda_2}{\lambda_1}=\frac{p}{q},\quad q>0,\quad p\leq 0,$$

其中p和q是两个整数，并且如果$p=0$，我们可令$\lambda_2=0$.

$$G_2=\left\{exp\left(\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}t\right)\,\middle|\, t\in\mathbb{R}\right\}. \tag{9.3.2}$$

$$G_3=\left\{exp\left(\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}t\right)\,\middle|\, t\in\mathbb{R}\right\}=SO(2,\mathbb{R}). \tag{9.3.3}$$

$$G_4=\left\{\prod_{i<\infty} exp(A_it_i)\,\middle|\, A_i=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\ \text{或}\ A_i=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix},\ t_i\in\mathbb{R}\right\}. \tag{9.3.4}$$

(2) 假设$n=2$时系统(9.1.1)关于$G=TG_iT^{-1}$状态空间对称，这里$T\in GL(2,\mathbb{R})$，那么(9.1.1)满足

$$f_j=\sum_{n=0}^{\infty}a_n^jp_n^i(T^{-1}y)TB_n^iT^{-1}y,\quad j=0,\cdots,m,\ i=1,2,3,4, \tag{9.3.5}$$

其中

$$p_n^1(x)=x_1^{-np}x_2^{nq};\qquad B_n^1=\begin{bmatrix}\alpha_n & 0\\ 0 & \beta_n\end{bmatrix};$$

$$p_n^2(x)=x_2^n;\qquad B_n^2=\begin{bmatrix}\alpha_n & \beta_n\\ 0 & \alpha_n\end{bmatrix};$$

$$p_n^3(x)=(x_1^2+x_2^2)^n;\quad B_n^3=\begin{bmatrix}\alpha_n & \beta_n\\ -\beta_n & \alpha_n\end{bmatrix};$$

$$p_n^4(x)=x_2^n;\qquad B_n^4=I.$$

证明 首先, 我们证明, 如果$n=2$ 时系统(9.1.1) 状态空间对称, 那么它可以表示成满足一定条件的特殊形式. 回顾(9.2.1), 容易看出(9.2.1) 具有形式

$$f_j(x)=\sum_{n=0}^{\infty}p_n^j(x)B_n^jx,\quad x\in\mathbb{R}^2,\quad j=0,\cdots,m. \tag{9.3.6}$$

(由于下面的讨论和j 无关, 为了记号的方便, 在下面的证明中我们省略j) 而且, 对任意$S\in o(2,\mathbb{R})$, 或者等价地选取一个基底

$$S=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix},$$

我们有

$$L_{Sx}p_n(x)=0, \tag{9.3.7}$$

$$[S,B_n]=0. \tag{9.3.8}$$

根据引理9.2.10, (9.1.1) 具有状态空间对称群$SO(2,\mathbb{R})$, 如果它具有(9.3.6) 的形式, 并且满足(9.3.7)~(9.3.8). □

为了完成证明, 我们需要下面的引理. 它说明上述形式和条件适用于所有平面状态空间对称系统.

引理 9.3.2 设$V\in gl(2,\mathbb{R})$, $G=\{e^{Vt}\,|\,t\in\mathbb{R}\}$. 一个平面系统

$$\dot{x}=f(x),\quad x\in\mathbb{R}^2 \tag{9.3.9}$$

关于G 对称, 当且仅当

(i) $f(x)$ 可以表示成(9.3.6);

(ii) p_n 和B_n 分别满足(9.3.7) 和(9.3.8).

证明 注意到ad_{Vx} 不改变$f(x)$ 中每个齐次分量的次数, 因此我们可以简单地假设$f(x)$ 是一个齐次向量场, 即令

$$f(x)=\begin{bmatrix}\sum\limits_{i=1}^{n}a_ix_1^{n-i}x_2^i\\ \sum\limits_{j=1}^{n}b_jx_1^{n-j}x_2^j\end{bmatrix}. \tag{9.3.10}$$

我们假设V 是Jordan 标准型.

情形1: 假设

$$V=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}.$$

利用引理9.1.5, $[Vx,f]=0$ 得到

$$\begin{cases}((n-i-1)\lambda_1+i\lambda_2)a_i=0,\\ ((n-j)\lambda_1+(j-1)\lambda_2)b_j=0, \quad i,j=0,\cdots,n.\end{cases} \tag{9.3.11}$$

为了得到非零的a_i, b_j, 我们需要

$$det\left(\begin{bmatrix}n-i-1 & i\\ n-j & j-1\end{bmatrix}\right)=(j-i-1)(n-1)=0. \tag{9.3.12}$$

如果$n=1$, 则$f(x)$ 是线性的, 直接计算可得结论成立. 我们考虑$n>1$ 的情形. 由(9.3.12) 我们有

$$j-i-1=0. \tag{9.3.13}$$

由(9.3.11) 我们有

$$(n-j)\lambda_1+(j-1)\lambda_2=0. \tag{9.3.14}$$

由于λ_1 和λ_2 不可能同时为0, 我们可以假设$\lambda_1\neq 0$, 并且令$\mu=\frac{\lambda_2}{\lambda_1}$. 根据(9.3.14), μ 是一个有理数. 首先, 我们假设$\lambda_2\neq 0$, 那么存在两个互质整数p, q, 使得

$$\mu=\frac{p}{q}. \tag{9.3.15}$$

于是由(9.3.14) 得到

$$\begin{cases}i=j-1=tq,\ q>0,\ p<0,\\ n=t(q-p)+1, \quad t=1,2,\cdots.\end{cases}$$

$f(x)$ 的形式为

$$f_{t(q-p)+1}(x)=x_1^{-tp}x_2^{tq}\begin{bmatrix}a_t & 0\\ 0 & b_t\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}, \quad t=1,2,\cdots. \tag{9.3.16}$$

另一方面, 考虑V-不变多项式. 假设

$$p_{n-1}(x)=\sum_{k=0}^{n-1}c_kx_1^{n-k-1}x_2^k.$$

由$L_{Vx}p_{n-1}(x)=0$ 我们有

$$(n-k-1)\lambda_1+k\lambda_2=0. \tag{9.3.17}$$

比较(9.3.17) 和(9.3.14) 可以看出$x_1^{-tp}x_2^{tq}$ 是(9.3.7) 在(λ_1,λ_2) 下的解的集合. 进一步假设$\lambda_1\neq\lambda_2$, 那么(9.3.16) 给出了满足(9.3.7)~(9.3.8) 的所有解.

现在假设$\lambda_2=0$, 容易看出满足(9.3.11) 的向量场具有如下形式

$$f_t(x)=x_2^{t-1}\begin{bmatrix}a_t & 0\\ 0 & b_t\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}, \tag{9.3.18}$$

它就是(9.3.7)~(9.3.8) 关于$\lambda_2=0$ 的解的集合.

最后, 假设λ_1,λ_2 是复数. 我们允许$f(x)$ 有复系数, 那么上面的论述仍然成立. 例如$\lambda_{1,2}=\alpha\pm\beta J$, 其中$J=\sqrt{-1}$, 那么由(9.3.14)~(9.3.15) 我们有$\alpha=0$, $\mu=-1$, 这样就有

$$V=\begin{bmatrix}J & 0\\ 0 & J\end{bmatrix}. \tag{9.3.19}$$

情形2: 假设

$$V=\begin{bmatrix}\lambda & 1\\ 0 & l\end{bmatrix}.$$

由引理9.1.5 得到

$$\begin{cases}\lambda(n-1)a_i+(n-i+1)a_{i-1}-b_i=0,\\ \lambda(n-1)b_i+(n-i+1)b_{i-1}=0, \quad i=0,\cdots,n+1,\end{cases} \tag{9.3.20}$$

其中为了记号的方便, 我们令$a_{-1}=b_{-1}=a_{n+1}=b_{n+1}=0$.

首先, 我们假设$\lambda\neq 0$, 利用(9.3.20) 的第二个方程, 并令$i=0$, 我们得到$b_0=0$, 递推下去, 我们可以得到所有的$b_i=0$. 于是由第一个方程得到所有的$a_i=0$. 因此没有非平凡解. 然后, 令$\lambda=0$, 第二个方程给出了非零解$b_n\neq 0$ 和$b_i=0$, $i\neq n$. 将它们代入第一个方程得到$a_n\neq 0$ 和$a_{n-1}=b_n\neq 0$, $a_i=0$, $i\leq n-2$, 于是f_n 的非平凡解变成

$$f_n=\begin{bmatrix}b_nx_1x_2^{n-1}+a_nx_2^n\\ b_nx_2^n\end{bmatrix}=x_2^{n-1}\begin{bmatrix}b_n & a_n\\ 0 & b_n\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}. \tag{9.3.21}$$

同理, 我们可以证明它由(9.3.7)~(9.3.8) 的所有的解组成.

最后, 我们考虑V 不是Jordan 标准型的情形. 作线性变换$y = Tx$, 方程(9.1.12) 变为

$$[T_*(Vx), T_*(f(x))] = 0.$$

现在假设$f(x)$ 具有(9.3.10) 的形式, 则

$$\begin{cases} T_*(Vx) = TVT^{-1}y, \\ T_*(f(x)) = \begin{bmatrix} \sum_{i=1}^{n} \tilde{a}_i y_1^{n-i} y_2^i, \\ \sum_{j=1}^{n} \tilde{b}_j y_1^{n-j} y_2^j \end{bmatrix}. \end{cases}$$

因此我们可以假设TVT^{-1} 具有Jordan 标准型. 假设它关于1 维群

$$G = \{e^{TVT^{-1}t} \,|\, t \in \mathbb{R}\}$$

对称, 那么原系统显然关于

$$G = \{e^{Vt} \,|\, t \in \mathbb{R}\}$$

对称. 因为(9.1.12) 是与坐标无关的, 而且在y 下系统具有(9.3.6) 的形式, 于是

$$T_*^{-1}(f_j(y)) = T_*^{-1}(p_n(y)B_n y) = p_n(Tx)T^{-1}B_n Tx,$$

即原系统也具有(9.2.15) 的形式. 由于(9.3.7) 和(9.3.8) 与坐标无关, 它们对原系统也成立. 证毕. □

下面的推广是引理9.3.2 的证明的一个直接结果.

引理 9.3.3 一个平面系统

$$\dot{x} = f(x), \quad x \in \mathbb{R}^2 \tag{9.3.22}$$

具有对称群$G < GL(2,\mathbb{R})$, 当且仅当

(i) $f(x)$ 可以表示成(9.3.6) 的形式;

(ii) p_n 和B_n 关于任意$S \in g(G)$ 满足(9.3.7) 和(9.3.8).

定理9.3.1 的证明(续) 下面我们考虑一个维数大于1 的可能的对称群G. 设$0 \neq A \in g(G)$, $g(G)$ 是G 的李代数.

情形1: 假设

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

由引理9.3.2 有

$$f = \sum_{i=0}^{\infty} c_i p_i(x) \begin{bmatrix} \alpha_i & 0 \\ 0 & b_i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

其中$p_i(x) = x_1^{-tp} x_2^{tq}$. 设

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \in g(G),$$

那么

$$\begin{aligned} L_{Bx} p_i(x) = & -tp x_1^{-tp-1} x_2^{tq} (b_{11} x_1 + b_{12} x_2) \\ & + tq x_1^{-tp} x_2^{tq-1} (b_{21} x_1 + b_{22} x_2) = 0. \end{aligned} \tag{9.3.23}$$

如果$p \neq 0$, 则有

$$\begin{cases} -b_{11} p + b_{22} q = 0, \\ b_{12} = b_{21} = 0, \end{cases}$$

于是

$$\frac{b_{22}}{b_{11}} = \frac{p}{q} = \frac{\lambda_2}{\lambda_1}.$$

即A 和B 是线性相关的, 且$\dim(G) = 1$. 为了探测新元素, 我们假设$p = 0$, 于是

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{等价地} \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

由(9.3.23) 我们有$b_{21} = b_{22} = 0$, 即

$$B = \begin{bmatrix} b_{11} & b_{12} \\ 0 & 0 \end{bmatrix}.$$

为了使

$$\left(B, \begin{bmatrix} a_i & 0 \\ 0 & b_i \end{bmatrix} \right) = 0,$$

显然如果$b_{12} \neq 0$, 则$a_i = b_i$. 我们得出结论

$$g = \operatorname{Span} \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}, \tag{9.3.24}$$

并且相应的系统是

$$\dot{x} = \sum_{i=0}^{\infty} a_i x_2^i \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{9.3.25}$$

在情形1 中, 如果$\lambda_2 \neq 0$, 我们可以交换x_1, x_2 来得到需要的形式. 实际上, 情形1, 2 和4 在上面已经讨论过, 只剩下情形3 没讨论. 在前面它被当作情形1 含有复特征值的特殊情形. 由情形1 中V 如(9.3.19) 所示的地方开始, 我们可以作如下变换: 令

$$x = Ty = \begin{bmatrix} 1 & 1 \\ J & -J \end{bmatrix}.$$

于是

$$T_*(Vy) = TVT^{-1}x = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x.$$

且

$$(T^{-1})^* p_{2n}(y) = p_{2n}(T^{-1}x) = \frac{a_n}{2^n}(x_1^2 + x_2^2)^n,$$

这就是所要的形式. □

注 如果系统(9.1.1) 满足(9.3.5), 那么直接验证可知, 它有相应的状态空间对称群TG_iT^{-1}. 因此定理9.3.1 对所有的平面状态空间对称系统和它们的状态空间对称群给出了一个完整的描述.

9.4 最大状态空间对称群

本节我们主要考虑对于一个给定的系统如何找出最大连通状态空间对称群. 首先我们需要做一些准备.

回忆第 5 章, 给定一个矩阵$A(x) \in M_{p\times q}$, 它的每个元素$a_{i,j}(x)$ ($x \in \mathbb{R}^n$) 都是光滑函数. 我们定义$A(x)$ 的微分$DA(x)$ 为一个$p \times nq$ 矩阵, 通过将$a_{i,j}$ 替换成它的微分$(\frac{\partial a_{i,j}}{\partial x_1}, \cdots, \frac{\partial a_{i,j}}{\partial x_n})$ 得到. 高阶微分可以递推定义为

$$D^{k+1}A(x) = D(D^k A(x)), \quad k \geq 1.$$

这个记号的优势是, 它对多元函数的表达极为方便, 例如, 对于多元函数的Taylor 级数展开, 可以得到它的系数为

$$f_k = \frac{1}{k!} D^k f(0), \quad k = 0, 1, \cdots.$$

利用换位矩阵$W_{[m,n]}$, 我们定义两个矩阵(具体见第5 章)

$$\Psi_k = \sum_{s=0}^{k} I_{n^s} \otimes W_{[n^{k-s},n]},$$

$$E_k^n := I_{n^{k-1}} \otimes W_{[n^{k-1},n]} \ltimes V_c(I_{n^{k-1}}).$$

于是我们有

定理 9.4.1 假设系统(9.1.1) 具有一个状态空间对称群G, 且G 的李代数为$g(G)$, 那么$\alpha \in g(G)$, 当且仅当$\xi = V_c(\alpha)$ 是如下代数方程的解.

$$\begin{aligned}\left(\left[T_N^{\mathrm{T}}(n,k) \otimes (f_k^j \Phi_{k-1})\right] E_k^n - \left[T_N^{\mathrm{T}}(n,k)(f_k^j)^{\mathrm{T}}\right] \otimes I_n\right)\xi = 0\\ k = 0,1,2,\cdots;\ j = 0,1,\cdots,m.\end{aligned} \tag{9.4.1}$$

证明 回顾(5.5.4)~(5.5.5), 我们有

$$D(x^{k+1}) = \Psi_k(x^k \otimes I_n) = \Psi_k \ltimes x^k. \tag{9.4.2}$$

利用(9.4.2), 容易证明如下公式:

$$\begin{aligned}L_{Vx} f_k x^k &= f_k \Psi_{k-1} x^{k-1} V x - V f_k x^k\\ &= f_k \Psi_{k-1}(I_{n^{k-1}} \otimes V)x^k - V f_k x^k, \quad k = 1,2,\cdots.\end{aligned} \tag{9.4.3}$$

为了得到唯一解, 我们将它转化到自然基底下, 于是可得

$$L_{Vx} f_k x^k = [f_k \Psi_{k-1}(I_{n^{k-1}} \otimes V) - V f_k] T_N(n,k) x_{(k)}, \quad k = 1,2,\cdots. \tag{9.4.4}$$

因此, 导数为0, 当且仅当

$$[f_k \Psi_{k-1}(I_{n^{k-1}} \otimes V) - V f_k] T_N(n,k) = 0, \quad k = 1,2,\cdots. \tag{9.4.5}$$

为了简化(9.4.5) 我们需要下面的公式(参见第3 章, 命题3.2.3)

$$V_c(AZB) = (B^{\mathrm{T}} \otimes A) V_c(Z). \tag{9.4.6}$$

再次利用(9.4.4), (9.4.5) 可以转化为

$$\begin{aligned}L_{Vx} f_k x^k &= (T_N^{\mathrm{T}}(n,k) \otimes (f_k \Psi_{k-1})) V_c(I_{n^{k-1}} \otimes V)\\ &\quad -((T_N^{\mathrm{T}}(n,k) f_k^{\mathrm{T}}) \otimes I_n) V_c(V) = 0, \quad k = 1,2,\cdots.\end{aligned} \tag{9.4.7}$$

□

为了将(9.4.7) 转化为一个标准的线性方程, 我们还需要两个公式, 当然这些公式本身也很重要.

命题 9.4.2 设$A \in M_{m\times n}$, $B \in M_{p\times q}$, 则

$$\begin{aligned} V_c(A\otimes B) &= (I_n \otimes W_{[m,q]}) \ltimes V_c(A) \ltimes V_c(B) \\ &= (I_n \otimes W_{[m,q]}) \ltimes W_{[pq,mn]} \ltimes V_c(B) \ltimes V_c(A). \end{aligned} \tag{9.4.8}$$

$$\begin{aligned} V_r(A\otimes B) &= W_{[mp,nq]}(I_n \otimes W_{[m,q]}) \ltimes W_{[m,n]} \ltimes (I_{mn} \otimes W_{p,q}) \\ &\quad \ltimes V_r(A) \ltimes V_r(B) \\ &= W_{[mp,nq]}(I_n \otimes W_{[m,q]}) \ltimes W_{[m,n]} \ltimes (I_{mn} \otimes W_{p,q}) \\ &\quad \ltimes W_{[pq,mn]} \ltimes V_r(B) \ltimes V_r(A). \end{aligned} \tag{9.4.9}$$

证明 我们只证明(9.4.8) 中的第一个公式, 其他的公式是它的直接结果. 首先, 我们假设$n = 1$, 于是显然有

$$V_c(A) \ltimes V_c(B) = \mathrm{col}(a_{11}B_1, \cdots, a_{11}B_q, \cdots, a_{m1}B_1, \cdots, a_{m1}B_q),$$

且

$$V_c(A\otimes B) = \mathrm{col}(a_{11}B_1, \cdots, a_{m1}B_1, \cdots, a_{11}B_q, \cdots, a_{m1}B_q).$$

注意它们由相同的但按照不同的双指标次序排列的p 维向量集合构成. 直接计算有

$$V_c(A\otimes B) = W_{[m,q]} \ltimes V_c(A) \ltimes V_c(B).$$

对于更一般的情况, 我们只需对n 个块作交换即可. 于是立即得到(9.4.8) 中的第一个公式成立. □

定理9.4.1 的证明(续) 记

$$E_k^n := I_{n^{k-1}} \otimes W_{[n^{k-1},n]} \ltimes V_c(I_{n^{k-1}}).$$

利用(9.4.8), 我们有

$$V_c(I_{n^{n-1}}) \otimes V = E \ltimes V_c(V).$$

将它代入到(9.4.7) 就得到定理9.4.1. □

定理9.4.1 给出了一个数值方法来计算系统(9.1.1) 的最大连通状态空间对称群.

例 9.4.3 考虑系统

$$\dot{x} = f(x) = f_3 x^3, \quad x \in \mathbb{R}^3, \tag{9.4.10}$$

其中$f_3 = (f_{ij}) \in M_{3\times 27}$. 设$C = (c_{ij}) \in M_{4\times 2}$ 是一个参数集合. 我们设系数矩阵为

$$\begin{aligned}
&f_{1,2} = f_{1,4} = f_{1,10} = c_{1,1},\\
&f_{1,3} = f_{1,7} = f_{1,19} = c_{1,2},\\
&f_{2,6} = f_{2,8} = f_{2,12} = f_{2,16} = f_{2,20} = f_{2,22} = c_{2,1};\\
&f_{3,6} = f_{3,8} = f_{3,12} = f_{3,16} = f_{3,20} = f_{3,22} = c_{2,2},\\
&f_{2,5} = f_{2,11} = f_{2,13} = c_{3,1},\\
&f_{3,5} = f_{3,11} = f_{3,13} = c_{3,2},\\
&f_{2,9} = f_{2,21} = f_{2,25} = c_{4,1},\\
&f_{3,9} = f_{3,21} = f_{3,25} = c_{4,2},\\
&f_{i,j} = 0, \quad \text{对其他 } (i,j).
\end{aligned}$$

仔细计算可知这样的参数群保证了非平凡对称群的存在性.

根据定理9.4.1, 我们可以构造矩阵

$$S_3^3 = T_N^{\mathrm{T}}(3,3) \otimes (f_3\Psi_2)E_2^3 - (T_N^{\mathrm{T}}(3,3)f_3^{\mathrm{T}}) \otimes I_3,$$

并且我们只需对$S_3^3\xi = 0$ 解出ξ.

情形1: 设C 是一个任意矩阵.

特别地, 如果我们选择$c_{12} = 2$, $c_{21} = c_{22} = c_{31} = c_{32} = 3$, 而其余的$c_{ij} = 1$, 那么通过计算可知$S_3^3$ 是一个30×9 的矩阵. 为了节省空间, 我们只列出它的非零元素:

$s_{1,2} = 3$	$s_{1,3} = 6$	$s_{4,1} = 3$	$s_{4,5} = 3$
$s_{4,6} = 6$	$s_{5,2} = 3$	$s_{5,3} = 18$	$s_{6,2} = 6$
$s_{6,3} = 15$	$s_{7,1} = 6$	$s_{7,8} = 3$	$s_{7,9} = 6$
$s_{8,2} = 12$	$s_{8,3} = 6$	$s_{9,2} = 18$	$s_{10,4} = 3$
$s_{10,7} = -3$	$s_{11,1} = 3$	$s_{11,5} = 3$	$s_{11,6} = 18$
$s_{11,8} = -3$	$s_{12,1} = 3$	$s_{12,5} = 6$	$s_{12,6} = 15$
$s_{12,9} = -3$	$s_{13,4} = -6$	$s_{13,7} = -12$	$s_{14,1} = 18$
$s_{14,6} = 6$	$s_{14,8} = -12$	$s_{14,9} = 18$	$s_{15,1} = 18$
$s_{15,5} = 18$	$s_{15,6} = -12$	$s_{15,8} = 6$	$s_{16,4} = -3$
$s_{16,7} = 9$	$s_{17,1} = 3$	$s_{17,5} = -3$	$s_{17,8} = 15$

$$
\begin{array}{llll}
s_{17,9}=6 & s_{18,1}=3 & s_{18,6}=-3 & s_{18,8}=18 \\
s_{18,9}=3 & s_{20,4}=3 & s_{21,4}=3 & s_{23,4}=18 \\
s_{23,7}=3 & s_{24,4}=18 & s_{24,7}=3 & s_{26,4}=3 \\
s_{26,7}=18 & s_{27,4}=3 & s_{27,7}=18 & s_{29,7}=3 \\
s_{30,7}=3 & & &
\end{array}
$$

非平凡解是

$$\xi=(1\,0\,0\,0\,-1\,0\,0\,0\,-1)^{\mathrm{T}}.$$

利用计算机可以验证对于任意的C, 这个ξ 总是解.

即含有上述参数的系统(9.4.1) 的最大连通不变线性群是

$$G_r=\left\{\exp\left(\begin{bmatrix}1&0&0\\0&-1&0\\0&0&-1\end{bmatrix}t\right)\middle|\, t\in\mathbb{R}\right\}.$$

情形2: 令$c_{ij}=1,\forall i,j$, 那么30×9 的矩阵S_3^3 具有如下的非零元素

$$
\begin{array}{llll}
s_{1,2}=3 & s_{1,3}=3 & s_{4,1}=3 & s_{4,5}=3 \\
s_{4,6}=3 & s_{5,2}=3 & s_{5,3}=6 & s_{6,2}=6 \\
s_{6,3}=3 & s_{7,1}=3 & s_{7,8}=3 & s_{7,9}=3 \\
s_{8,2}=3 & s_{8,3}=6 & s_{9,2}=6 & s_{9,3}=3 \\
s_{10,4}=3 & s_{10,7}=-3 & s_{11,1}=3 & s_{11,5}=3 \\
s_{11,6}=6 & s_{11,8}=-3 & s_{12,1}=3 & s_{12,5}=6 \\
s_{12,6}=3 & s_{12,9}=-3 & s_{14,1}=6 & s_{14,6}=6 \\
s_{14,9}=6 & s_{15,1}=6 & s_{15,5}=6 & s_{15,8}=6 \\
s_{16,4}=-3 & s_{16,7}=3 & s_{17,1}=3 & s_{17,5}=-3 \\
s_{17,8}=3 & s_{17,9}=6 & s_{18,1}=3 & s_{18,6}=-3 \\
s_{18,8}=6 & s_{18,9}=3 & s_{20,4}=3 & s_{21,4}=3 \\
s_{23,4}=6 & s_{23,7}=3 & s_{24,4}=6 & s_{24,7}=3 \\
s_{26,4}=3 & s_{26,7}=6 & s_{27,4}=3 & s_{27,7}=6 \\
s_{29,7}=3 & s_{30,7}=3 & &
\end{array}
$$

解是

$$
\begin{cases}
\xi_1=(0\,0\,0\,0\,-1\,1\,0\,1\,-1)^{\mathrm{T}},\\
\xi_2=(-2\,0\,0\,0\,1\,1\,0\,1\,1)^{\mathrm{T}}.
\end{cases}
$$

当我们将ξ_1 和ξ_2 转换回矩阵, 仍记它们为$\xi_1,\xi_2 \in gl(n,\mathbb{R})$, 那么$[\xi_1,\xi_2]=0$. 即它们是可交换的, 这就意味着$g=\mathrm{Span}\{\xi_1,\xi_2\}$ 是一个李代数. 于是对于这组系数, 系统(9.4.1) 的最大连通不变线性群是

$$G_1=\left\{\exp\left(\begin{bmatrix}0&0&0\\0&-1&1\\0&1&-1\end{bmatrix}t_1\right)\exp\left(\begin{bmatrix}-2&0&0\\0&1&1\\0&1&1\end{bmatrix}t_2\right)\middle|\, t_1,t_2\in\mathbb{R}\right\}.$$

我们可以再研究一下情形2. 仔细计算有

$$\begin{aligned}f(x) &= f_3x^3=\begin{bmatrix}3x_1^2(x_2+x_3)\\3x_1x_2^2+6x_1x_2x_3+3x_1x_3^2\\3x_1x_2^2+6x_1x_2x_3+3x_1x_3^2\end{bmatrix}\\&= 3x_1(x_2+x_3)\begin{bmatrix}1&0&0\\0&1&1\\0&1&1\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}:=p(x)Kx.\end{aligned}$$

于是容易验证系统满足(9.3.7)~(9.3.8), 即$p(x)$ 是$g(G_1)$ 不变的, 且

$$\begin{bmatrix}1&0&0\\0&1&1\\0&1&1\end{bmatrix}\in Z(g(G_1)).$$

□

9.5 对称性和能控性

本节我们主要讨论具有对称性的仿射非线性系统的一些能控性问题.

首先, 我们考虑线性系统

$$\dot{x}=Ax+\sum_{i=1}^{m}b_iu_i:=Ax+Bu,\quad x\in\mathbb{R}^n,\ u\in\mathbb{R}^m. \tag{9.5.1}$$

它的线性对称性可以描述为下面的命题.

命题 9.5.1 系统(9.5.1) 具有一个连通状态空间对称群$G<GL(n,\mathbb{R})$, 当且仅当对于任意$\alpha\in g(G)$ 有

$$\begin{cases}\alpha A-A\alpha=0,\\ \alpha B=0.\end{cases} \tag{9.5.2}$$

证明 必要性. 由于$\alpha \in g(G)$, $e^{\alpha t} \in G$, $\forall t \in \mathbb{R}$. 又由于系统关于G 对称, 根据定义我们有

$$(e^{\alpha t})_*(Ax + Bu) = e^{\alpha t}Ae^{-\alpha t}x + e^{\alpha t}B = Ax + Bu, \quad \forall x \in \mathbb{R}^n, \ u, v \in \mathbb{R}^m.$$

因此

$$\begin{cases} e^{\alpha t}Ae^{-\alpha t} = A, \\ e^{\alpha t}B = B. \end{cases} \tag{9.5.3}$$

对(9.5.3) 的第一个方程两边关于t 进行微分, 我们有

$$\alpha e^{\alpha t}Ae^{-\alpha t} - e^{\alpha t}A\alpha e^{-\alpha t} = 0.$$

令$t = 0$ 得到(9.5.2) 的第一个方程. 同理, 我们可以得到第二个.

充分性. 利用$e^{\alpha t}$ 的Taylor 级数展开, 可以看出由(9.5.2) 得到(9.5.3). 由(9.1.13) 中的G 的结构得出结论成立. □

将(9.5.2) 表示成矩阵形式, 我们有

推论 9.5.2 系统(9.5.1) 具有一个非平凡状态空间对称群($G \neq \{I_n\}$), 当且仅当方程

$$\begin{bmatrix} A^{\mathrm{T}} \otimes I_n - I_n \otimes A \\ B^{\mathrm{T}} \otimes I_n \end{bmatrix} \xi = 0 \tag{9.5.4}$$

具有一个非零解.

推论 9.5.3 如果系统(9.5.1) 是完全能控的, 那么它没有非平凡线性状态空间对称群$G < GL(n, \mathbb{R})$.

证明 假设$\Phi \in G$, 则

$$\Phi A \Phi^{-1} = A, \quad \Phi B = B.$$

因此

$$\Phi(A^{n-1}B, \cdots, B) = (A^{n-1}B, \cdots, B),$$

这样$\Phi = I_n$. □

现在假设系统(9.1.1) 具有一个状态空间对称群G (G 可能不是线性群) . 对于任意$\alpha \in G$, 映射$\theta_\alpha : \mathbb{R}^n \to \mathbb{R}^n$ 是一个微分同胚. 记

$$J_\alpha = \frac{\partial \theta_\alpha(x)}{\partial x}(0), \quad \alpha \in G.$$

利用系统在θ_α 上的Taylor 级数展开和验证线性项, 可以证明下面的结果:

命题 9.5.4 假设$f_0(0)=0$, 且系统(9.1.1) 具有一个状态空间对称群G, 则
(1) $G_L := \{ J_\alpha \mid \alpha \in G\} < GL(n,\mathbb{R})$ 是一个李子群;
(2) 设$A=\frac{\partial f_0}{\partial x}(0)$, $b_i=f_i(0)$, $i=1,\cdots,m$, 则线性近似系统

$$\dot{z}=Az+\sum_{i=1}^{m} b_i u_i$$

具有状态空间对称群G_L.

下面的结果可以看作对于一般对称的一个必要条件(其中G_L, A, B 如命题9.5.4 中所述).

推论 9.5.5 假设$f_0(0)=0$, 系统(9.1.1) 具有一个状态空间对称群G, 并且(A,B) 能控, 则

$$G_L=\{I_n\}.$$

考虑非线性控制系统(9.1.1). 定义其能控性李代数为[68]

$$\mathcal{L}=\{f_0,f_1,\cdots,f_m\}_{LA}, \tag{9.5.5}$$

则$\mathcal{L}(x)$ 可以看作状态空间$\mathbb{R}^n$ 上的一个分布. 如果这个分布在x 点维数为n, 则称系统在该点满足能控性秩条件. 它是非线性系统能控性的一个基本判定指标.

对称性可能引起不可控性[89], 下面是一个这样的结论:

命题 9.5.6 假设系统(9.1.1) 具有一个非平凡状态空间对称群$G<GL(n,\mathbb{R})$, 那么它在原点处不满足能控性秩条件.

证明 由于G 是非平凡的, 这就意味着存在$0\neq V\in g(G)$. 利用引理9.1.5, 我们有

$$[Vx, f_i(x)]=0, \quad i=0,1,\cdots,m.$$

利用Jacobi 恒等式, 对于能控性李代数

$$\mathcal{L}=\{f_0,f_1,\cdots,f_m\}_{LA},$$

我们也有

$$[Vx,\mathcal{L}]=0.$$

如果$\dim \mathrm{Span}\{\mathcal{L}\}(0)=n$, 我们可以找到在$x=0$ 处线性无关的$\eta_1(x),\cdots,\eta_n(x)\in\mathcal{L}$. $\eta_i(x)$ 可以表示成

$$\eta_i(x)=\eta_i(0)+O(\|x\|), \quad i=1,\cdots,n.$$

于是

$$0 = [Vx, \eta_i(x)] = -V\eta_i(0) + O(\|x\|), \quad i = 1, \cdots, n,$$

这就意味着

$$V\eta_i(0) = 0, \quad i = 1, \cdots, n.$$

因此, $V = 0$, 这就导致了矛盾. □

注意, 实际上上述命题说系统(9.1.1) 在任意$x_0 \in \mathbb{R}^n$ 处不满足能控性秩条件, 如果它是在任意点x_0 处关于非平凡$G < GL(n,\mathbb{R})$ 状态空间对称. 这里"在点x_0 处对称"的意思是, 对于任意的$\alpha \in G$, 系统在作用$\theta_\alpha : x - x_0 \mapsto \alpha(x - x_0)$ 下不变.

9.6　注释与参考

自然界充满了各种对称性, 而任何一种对称都可以看作对象在某种群作用下的不变性[103]. 因此, 研究对称就是研究群作用下的不变性.

关于动力系统对称性方面的专著很多, 例如文献[7], [81], [83], [84].

关于控制系统的对称性也已经有很多学者研究. 这方面的部分结果可见文献[55], [114], [127], [88], [89].

"状态空间对称" 和"系统对称" 是控制系统对称性中较有代表性的两种对称性. 前者在文献[55], [127] 中提出和使用, 而后者则来自于文献[89] . 容易看出状态空间对称是系统对称的一个特殊情形. 本章讨论的主要是状态空间对称中的一个特殊情况: 线性状态空间对称, 是指状态空间在一般线性群$GL(n,\mathbb{R})$ 的某个子群作用下的对称性.

本章的主要内容是基于文献[34] 中的工作改写的.

第 10 章　动态系统的稳定域

研究一个动态系统的稳定平衡点的稳定域 (也称吸引域) 是很重要的, 因为它可以决定一个工程系统(例如机械系统、化学反应堆等) 的安全工作区域. 特别地, 这个问题在电力系统中被广泛研究, 因为这里存在着多平衡点. 本章利用半张量积给出计算稳定域边界的公式.

10.1　稳定域的描述

在动态系统稳定性分析中, 例如在电力系统的紧急控制设计中, 吸引域和它的边界是在紧急突变过程稳定性计算中我们所关心的基本问题. 关键是突变的过程轨线是否会冲出变化后的稳态系统平衡点的稳定域. 只有这些可能的突变过程轨线在稳态平衡点的稳定域内, 系统才是安全的. 这个问题是电力系统安全稳定运行的关键.

近年来常用的方法是所谓的能量函数法, 它的基本思想是使用经过临界点(不稳定平衡点) 的能量函数的等值曲面来近似稳定边界.

为了叙述更准确, 我们需要给出对这个问题进行严格几何描述的一些记号和定义.

考虑如下形式的光滑非线性系统

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n, \tag{10.1.1}$$

其中$f(x)$ 是一个解析向量场.

设x_e 是(10.1.1) 的一个平衡点. x_e 的稳定和不稳定子流形, 分别记作$W^s(x_e)$ 和$W^u(x_e)$, 定义如下:

$$\begin{aligned} W^s(x_e) &= \left\{ p \in \mathbb{R}^n \,\middle|\, \lim_{t\to\infty} x(t,p) \to x_e \right\}, \\ W^u(x_e) &= \left\{ p \in \mathbb{R}^n \,\middle|\, \lim_{t\to-\infty} x(t,p) \to x_e \right\}. \end{aligned} \tag{10.1.2}$$

设x_s 是(10.1.1) 的一个稳定平衡点. x_s 的吸引域定义为

$$A(x_s) = \left\{ p \in \mathbb{R}^n \,\middle|\, \lim_{t\to\infty} x(t,p) \to x_s \right\}. \tag{10.1.3}$$

吸引域的边界记为$\partial A(x_s)$.

称一个平衡点是双曲的, 如果f 在x_e 处的Jacobi 矩阵$J_f(x_e)$ 没有零实部特征值. 称一个双曲平衡点是k 型的, 如果$J_f(x_e)$ 有k 个正实部特征值.

文献[122] 和[35] 证明了下述结论: 对于一个动力系统的稳定平衡点x_s, 如果它满足下述三个假设条件, 则稳定域边界由边界上的不稳定平衡点的稳定子流形构成. 见图10.1.1.

(i) 稳定边界$\partial A(x_s)$ 上的平衡点是双曲的;

(ii) 稳定边界$\partial A(x_s)$ 上的平衡点的稳定和不稳定子流形满足横截条件;

(iii) 稳定边界$\partial A(x_s)$ 上的每条轨线当$t \to \infty$ 时趋向于一个平衡点.

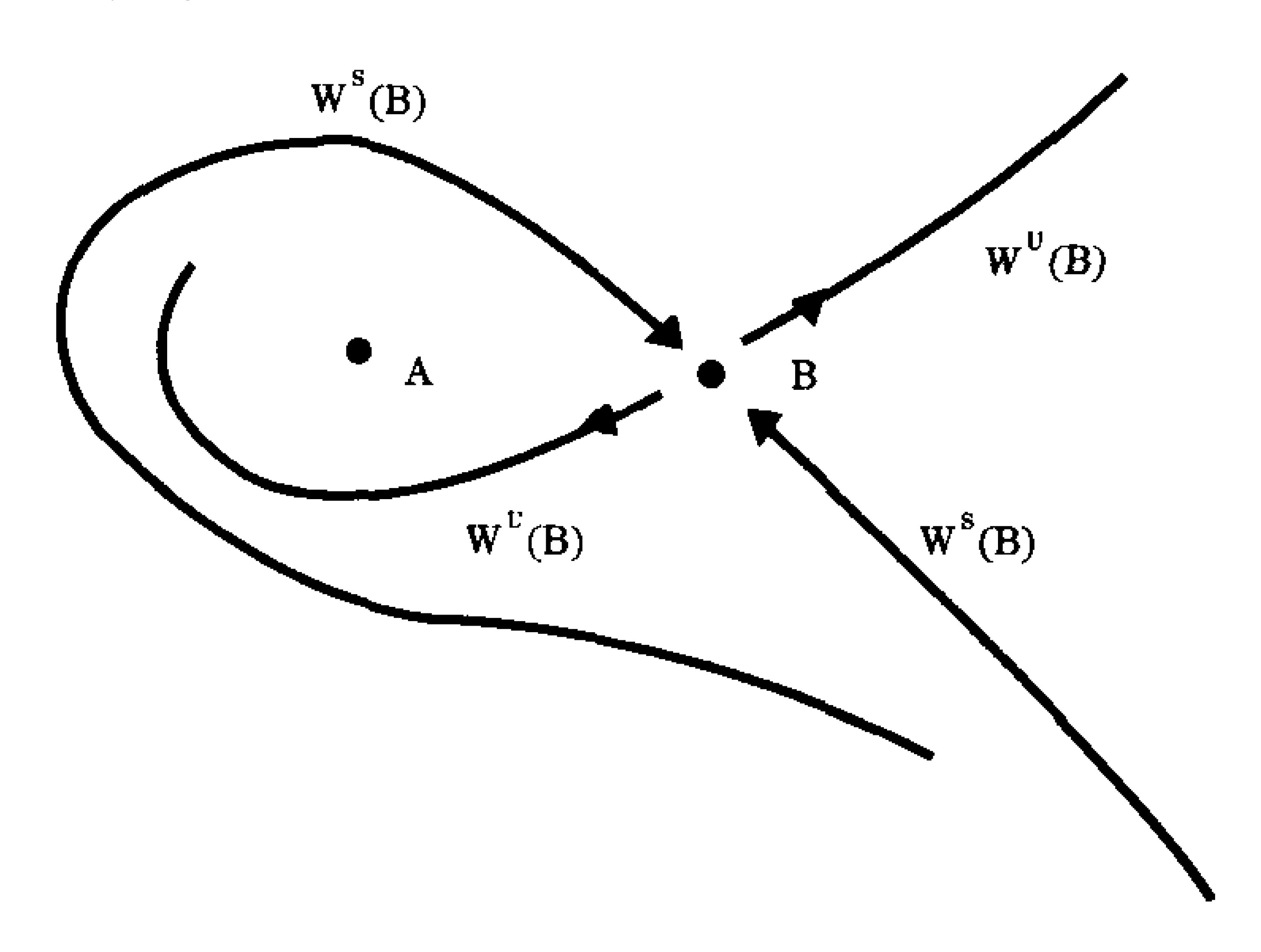

图 10.1.1 稳定域边界

一个微分流形M 的两个子流形N, S 称为是满足横截条件的, 如果对于任一交点$x \in N \cap S$, 两个子流形N, S 在该点的切空间的并等于流形在该点的切空间, 即

$$T_x(N) \cup T_x(S) = T_x(M).$$

众所周知, 如果状态空间是n 维的, 则稳定边界是$n-1$ 维的[35]. 因此, 稳定边界是由边界上的1 型平衡点的稳定子流形的闭包组成. 基于这个事实, 计算或估计1 型平衡点的稳定子流形具有重要的意义. 近来在电力系统的研究中除能量函数方法外, 还有其他许多算法来得到1 型平衡点的近似.

本章我们将给出稳定子流形的Taylor 展开公式. 它的前两项形成了最好的二次近似. 这个公式基于左半张量积.

10.2 稳定子流形方程

本节我们将考虑1 型平衡点的稳定子流形的展开.

不失一般性, 我们可以假设1 型平衡点是$x_u = 0$. 将(10.1.1) 中向量场f 由Taylor 展式展开成

$$f(x) = \sum_{i=1}^{\infty} F_i x^i = Jx + F_2 x^2 + \cdots, \tag{10.2.1}$$

其中$F_1 = J = J_f(0)$, $F_i = \frac{1}{i!}D^i f(0)$ 是已知的$n \times n^i$ 矩阵.

我们用$A^{-\mathrm{T}}$ 表示A 的转置A^{T} 的逆. 矩阵A 称为一个双曲矩阵, 如果它没有零实部的特征值.

引理 10.2.1 设A 是一个双曲矩阵. V_s 和V_u 分别记为A 的稳定和不稳定子空间, U_s 和U_u 分别记为$A^{-\mathrm{T}}$ 的稳定和不稳定子空间, 那么

$$V_s^{\perp} = U_u, \quad V_u^{\perp} = U_s. \tag{10.2.2}$$

证明 设A 是k 型的, 那么我们可以将A 转化为Jordan 标准型

$$Q^{-1}AQ = \begin{bmatrix} J_s & 0 \\ 0 & J_u \end{bmatrix},$$

其中J_s 和J_u 分别表示稳定和不稳定块. 分割$Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}$, 其中$Q_1$ 和Q_2 分别是前$n-k$ 列和后k 列, 那么

$$V_s = \operatorname{Span} col\{Q_1\}, \quad V_u = \operatorname{Span} col\{Q_2\}.$$

容易看出

$$Q^{\mathrm{T}}A^{-\mathrm{T}}Q^{-\mathrm{T}} = \begin{bmatrix} J_s^{-\mathrm{T}} & 0 \\ 0 & J_u^{-\mathrm{T}} \end{bmatrix}.$$

分割$Q^{-\mathrm{T}} = \begin{bmatrix} \tilde{Q}_1 & \tilde{Q}_2 \end{bmatrix}$, 其中$\tilde{Q}_1$ 和$\tilde{Q}_2$ 分别是$Q^{-\mathrm{T}}$ 的前$n-k$ 列和后k 列, 那么

$$U_s = \operatorname{Span} col\{\tilde{Q}_1\}, \quad U_u = \operatorname{Span} col\{\tilde{Q}_2\}.$$

于是由$Q^{-1}Q = I$ 得出结论. □

下面的推论是上面引理的一个直接结果.

推论 10.2.2 设A 是一个1 型矩阵, 且它的唯一一个不稳定特征值为μ. 设η 是A^{T} 关于特征值μ 的特征向量, 那么η 垂直于A 的稳定子空间.

证明 因为$A^{-\mathrm{T}}$ 的唯一一个不稳定特征值是$\frac{1}{\mu}$, 记η 是$A^{-\mathrm{T}}$ 的特征向量, 于是根据引理10.2.1 中的$\operatorname{Span}\{\eta\} = U_u = V_s^{\perp}$, 我们只需证明$\eta$ 也是A^{T} 关于μ 的特征向量即可.

由于

$$A^{-T}\eta = \frac{1}{\mu}\eta \Rightarrow A^{\mathrm{T}}\eta = \mu\eta,$$

于是结论成立. □

不失一般性, 我们假设后面的$x_u = 0$ 是系统(10.1.1) 的1 型平衡点.

下面的定理给出了一组1 型平衡点的稳定子流形所要满足的充分必要条件.

定理 10.2.3　设$x_u = 0$ 是系统(10.1.1) 的1 型平衡点.

$$W^s(e_u) = \{x \,|\, h(x) = 0\}. \tag{10.2.3}$$

那么$h(x)$ 是由下面的充要条件(10.2.4)~(10.2.6) 唯一决定的.

$$h(0) = 0, \tag{10.2.4}$$

$$h(x) = \eta^{\mathrm{T}} x + O(\|x\|^2), \tag{10.2.5}$$

$$L_f h(x) = \mu h(x), \tag{10.2.6}$$

其中$L_f h(x)$ 是$h(x)$ 关于f 的李导数, η 是$J_f^{\mathrm{T}}(0)$ 关于它的唯一一个正特征值μ 的特征向量.

证明　必要性. (10.2.4) 和(10.2.5) 的必要性显然. 我们只需证明(10.2.6) 的必要性. 首先, 注意到

$$\frac{\partial h}{\partial x} = \eta^{\mathrm{T}} + O(\|x\|). \tag{10.2.7}$$

因此局部存在零点的一个邻域, 使得

$$\mathrm{rank}(h(x)) = 1, \quad x \in U. \tag{10.2.8}$$

由于$W^s(e_u)$ 是f 不变的, 于是

$$\begin{cases} h(x) = 0, \\ L_f h(x) = 0, \quad x \in W^s(e_u). \end{cases} \tag{10.2.9}$$

由于$\dim(W^s(e_u)) = n - 1$, 得到

$$\mathrm{rank}\left(\begin{bmatrix} h(x) \\ L_f h(x) \end{bmatrix}\right) = 1,$$

这就说明$h(x)$ 和$L_f h(x)$ 应是线性相关的. 直接计算得

$$L_f h(x) = \eta^{\mathrm{T}} J_f(0) x + O(\|x\|^2) = \mu \eta^{\mathrm{T}} x + O(\|x\|^2).$$

于是对于$x \in U$, 由$h(x)$ 和$L_f h(x)$ 的线性相关性得到(10.2.6). 最后, 由系统的解析性, (10.2.6) 在全局也是对的.

充分性. 首先, 我们证明如果$h(x)$ 满足(10.2.4)~(10.2.6), 那么局部地

$$\{x \in U \mid h(x) = 0\}$$

在U 上是稳定子流形. 根据秩条件(10.2.8), 我们知道(参见文献[10], 定理5.8)

$$V := \{x \in U \mid h(x) = 0\}$$

是一个$n-1$ 维流形.

其次, 由于$L_f h(x) = 0$, V 是局部f 不变的. 最后, (10.2.5) 表明在局部零点是f 限制在V 上的渐近稳定平衡点. 因此, 在局部V 是(10.1.1) 的稳定子流形. 但是稳定子流形是唯一的[14], 因此V 就是$W^s(e_u)$.

由于系统是解析的, $\{x \mid h(x) = 0\}$ 与$W^s(e_u)$ 全局重合. □

10.3　二次近似

本节介绍不稳定平衡点的稳定子流形的二次近似. 与已知的二次近似[101, 91]相比, 本方法主要有两个优点: 1. 给出了精确的公式; 2. 它是唯一的误差仅为$O(\|x\|^3)$ 的公式.

为了计算方便, 我们记$h(x)$ 的Taylor 展开为

$$h(x) = H_1 x + H_2 x^2 + H_3 x^3 + \cdots = H_1 x + \frac{1}{2} x^{\mathrm{T}} \Psi x + H_3 x^3 + \cdots . \tag{10.3.1}$$

注意到在上面我们使用的二次项的两种形式: 半张量积形式$H_2 x^2$ 和二次型形式$\frac{1}{2}x^{\mathrm{T}}\Psi x$, 其中$\Psi = \mathrm{Hess}(h(0))$ 是$h(x)$ 在$x = 0$ 处的Hessian 矩阵, $H_2 = V_c^{\mathrm{T}}(\frac{1}{2}\Psi)$ 是$\frac{1}{2}\Psi$ 的列展开.

注意: 设$f(x, y)$ 为$\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ 的一个实值函数, 则它的Hessian 矩阵为

$$\mathrm{Hess}(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial y_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial y_m} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial y_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial y_m} \end{bmatrix}.$$

引理 10.3.1　(10.3.1) 的稳定子流形的二次型形式满足

$$\Psi\left(\frac{\mu}{2}I - J\right) + \left(\frac{\mu}{2}I - J^{\mathrm{T}}\right)\Psi = \sum_{i=1}^{n} \eta_i \,\mathrm{Hess}(f_i(0)), \tag{10.3.2}$$

其中μ 和η 依推论10.2.2 中的定义, $\mathrm{Hess}(f_i)$ 是f 的第i 个分量的Hessian 矩阵.

证明 首先, $h(x)=0$ 的线性近似是

$$H_1x=0,$$

它就是$W^s(x_u)$ 的稳定子流形的切空间. 由于η 是$W^s(x_u)$ 在x_u 处的法线方向, 显然$H_1=\eta$.

根据定理10.2.3, 李导数

$$L_fh(x)=0.$$

利用(5.5.4), 我们有

$$Dh(x)=H_1+H_2\Phi_1x+H_3\Phi_2x^2+\cdots=H_1+x^{\mathrm{T}}\Psi+H_3\Phi_2x^2+\cdots.$$

注意到向量场f 可以表示成

$$f(x)=Jx+\frac{1}{2}\begin{bmatrix}x^{\mathrm{T}}\operatorname{Hess}(f_1(0))x\\ \vdots\\ x^{\mathrm{T}}\operatorname{Hess}(f_n(0))x\end{bmatrix}+O(\|x\|^3).$$

计算L_fh 得到

$$\begin{aligned}L_fh &= \eta^{\mathrm{T}}Jx+x^{\mathrm{T}}\left(\frac{1}{2}\sum_{i=1}^{n}\eta_i\operatorname{Hess}(f_i(0))+\Psi J\right)x+O(\|x\|^3)\\ &= \mu\eta^{\mathrm{T}}x+x^{\mathrm{T}}\left(\frac{1}{2}\sum_{i=1}^{n}\eta_i\operatorname{Hess}(f_i(0))+\Psi J\right)x+O(\|x\|^3).\end{aligned}\tag{10.3.3}$$

注意到作为f 的不变子流形, 我们有

$$W^s(e_u)=\{x\mid h(x)=0,\ L_fh(x)=0\}.\tag{10.3.4}$$

利用(10.3.1) 和(10.3.4), 对于$W^s(e_u)$ 我们有

$$x^{\mathrm{T}}\left(\frac{1}{2}\sum_{i=1}^{n}\eta_i\operatorname{Hess}(f_i(0))+\Psi(J-\frac{\mu}{2}I)\right)x+O(\|x\|^3)=0.\tag{10.3.5}$$

将二次型形式表示成对称形式, 我们有(10.3.2). □

引理 10.3.2 (10.3.2) 具有唯一对称解.

证明 将(10.3.2) 表示成线性方程形式, 我们有

$$(A\otimes I_n+I_n\otimes A)V_c(\Psi)=V_c\left(\sum_{i=1}^{n}\eta_i\operatorname{Hess}(f_i(0))\right),\tag{10.3.6}$$

其中

$$A=\frac{\mu}{2}I-J^{\mathrm{T}}.$$

(10.3.6) 的形式可由标准Lyapunov 映射得到(参见文献[18]) . 众所周知(参见第3章), 设$\lambda_i\in\sigma(A)$, $i=1,\cdots,n$ 是A 的特征值, 则$A\otimes I_n+I_n\otimes A$ 的特征值是

$$\{\lambda_i+\lambda_j\,|\,1\le i,j\le n,\ \lambda_t\in\sigma(A)\}.$$

为了证明$A\otimes I_n+I_n\otimes A^{\mathrm{T}}$ 是非奇异的, 我们只需证明所有的$\lambda_i+\lambda_j\neq 0$. 设$\xi_i\in\sigma(J)$, $i=1,\cdots,n$ 是J 的特征值, 则

$$\lambda_i=\frac{\mu}{2}-\xi_i,\quad i=1,\cdots,n.$$

观察J 的特征值可以看出A 的唯一负特征值是$-\frac{\mu}{2}$, 并且A 的所有其他特征值的正实部都大于$\frac{\mu}{2}$, 于是

$$\lambda_i+\lambda_j\neq 0,\quad 1\le i,\ j\le n.$$

因此(10.3.2) 具有唯一解. 最后, 证明解是对称的. 可以验证

$$(A\otimes I_n+I_n\otimes A)W_{[n]}=W_{[n]}(A\otimes I_n+I_n\otimes A).\tag{10.3.7}$$

利用(10.3.7), 我们有

$$\begin{aligned}&(A\otimes I_n+I_n\otimes A)V_r(\Psi)=(A\otimes I_n+I_n\otimes A)W_{[n]}V_c(\Psi)\\&=W_{[n]}(A\otimes I_n+I_n\otimes A)V_c(\Psi)=W_{[n]}V_c\left(\sum_{i=1}^n\eta_i\,\mathrm{Hess}(f_i(0))\right)\\&=V_r\left(\sum_{i=1}^n\eta_i\,\mathrm{Hess}(f_i(0))\right)=V_c\left(\sum_{i=1}^n\eta_i\,\mathrm{Hess}(f_i(0))\right).\end{aligned}\tag{10.3.8}$$

最后一个等号可由$\sum\limits_{i=1}^n\xi_i\,\mathrm{Hess}(f_i(0))$ 是对称矩阵得到, 因此它的行展开和列展开相同. (10.3.8) 表明$V_r(\Psi)$ 是(10.3.6) 的另一个解, 但是(10.3.6) 仅有一个解, 因此我们有

$$V_r(\Psi)=V_c(\Psi).$$

即Ψ 对称. □

记V_c^{-1} 是V_c 的逆映射, 即它将矩阵A 从它的行展开$V_c(A)$ 还原为A.

综合引理10.2.1, 10.3.1, 10.3.2, 我们有下面的稳定子流形的二次形式的近似.

定理 10.3.3 x_u 的稳定子流形$h(x)=0$ 可以表示成

$$h(x) = H_1 x + \frac{1}{2} x^{\mathrm{T}} \Psi x + O(\|x\|^3), \tag{10.3.9}$$

其中

$$\begin{cases} H_1 = \eta^{\mathrm{T}}, \\ \Psi = V_c^{-1} \left\{ \left[\left(\frac{\mu}{2} I_n - J^{\mathrm{T}}\right) \otimes I_n + I_n \otimes \left(\frac{\mu}{2} I_n - J^{\mathrm{T}}\right) \right]^{-1} V_c \left(\sum\limits_{i=1}^{n} \eta_i \,\mathrm{Hess}(f_i(0)) \right) \right\}, \end{cases}$$

μ 和η 如推论10.2.2 中对$J=F_1$ 的定义, $\mathrm{Hess}(f_i)$ 是f 的第i 个分量f_i 的Hessian 矩阵.

注 如果e_u 是$n-1$ 型的, μ 是唯一的负特征值, 它的特征向量是η, 则上面的论述仍可适用于不稳定子流形. 特别地, (10.3.9) 是不稳定子流形的二次近似.

观察(10.3.5), 下面的推论是一个直接结果, 它可以简化一些计算.

推论 10.3.4 设

$$\sum_{i=1}^{n} \eta_i \,\mathrm{Hess}(f_i(0)) \left(\frac{\mu}{2} I_n - J \right)^{-1}$$

对称, 那么稳定子流形的二次近似是

$$h(x) = \eta^{\mathrm{T}} x + \frac{1}{4} x^{\mathrm{T}} \sum_{i=1}^{n} \eta_i \,\mathrm{Hess}(f_i(0)) \left(\frac{\mu}{2} I_n - J \right)^{-1} x = 0. \tag{10.3.10}$$

例 10.3.5 考虑系统

$$\begin{cases} \dot{x}_1 = x_1, \\ \dot{x}_2 = -x_2 + x_1^2, \quad x \in \mathbb{R}^2. \end{cases} \tag{10.3.11}$$

它的稳定和不稳定子流形是[91]

$$W^s(0) = \{x \in \mathbb{R}^2 \mid x_1 = 0\},$$

$$W^u(0) = \{x \in \mathbb{R}^2 \mid x_2 = \tfrac{1}{3} x_1^2\}.$$

我们用它们来验证公式(10.3.10). 对于(10.3.11), 我们有

$$J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

对于稳定子流形$W^s(0)$, 容易验证稳定特征值是$\mu = 1$, 特征向量是$\eta = (1\ 0)^{\mathrm{T}}$, 且

$$\mathrm{Hess}(f_1(0)) = 0, \quad \mathrm{Hess}(f_2(0)) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

于是

$$\frac{1}{4}\sum_{i=1}^{2} \eta_i \,\mathrm{Hess}(f_i(0)) \left(\frac{1}{2}I - J\right)^{-1} = 0,$$

即

$$h_s(x) = (1\ 0)x + 0 + O(\|x\|^3) = x_1 + O(\|x\|^3).$$

对于不稳定子流形$W^u(0)$, 易验证不稳定特征值是$\mu = -1$, 特征向量是$\eta = (0\ 1)^{\mathrm{T}}$, 于是

$$\frac{1}{4}\sum_{i=1}^{2} \eta_i \,\mathrm{Hess}(f_i(0))(\frac{-1}{2}I - J)^{-1} = \begin{bmatrix} -\frac{1}{3} & 0 \\ 0 & 0 \end{bmatrix}.$$

即

$$h_u(x) = (0\ 1)x + x^{\mathrm{T}} \begin{bmatrix} -\frac{1}{3} & 0 \\ 0 & 0 \end{bmatrix} x + O(\|x\|^3) = x_2 - \frac{1}{3}x_1^2 + O(\|x\|^3).$$

注意到我们用下一节的结论可以看出$h_s(x)$ 和$h_u(x)$ 的误差项$O(\|x\|^3)$ 都是0. 另外, 利用定理10.2.3 也可直接验证. 例如, 我们验证$h_u(x)$: 假设$h_u(x) = x_2 - \frac{1}{3}x_1^2$, 于是$W^u(e_u) = \{x \mid h_u(x) = 0\}$ 当且仅当由$h_u(x) = 0$ 得到$L_f h_u(x) = 0$, 而这是对的, 因为

$$L_f(h_u(x)) = \begin{bmatrix} -\frac{2}{3}x_1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ -x_2 + x_1^2 \end{bmatrix} = -x_2 + \frac{1}{3}x_1^2 = -h_u(x).$$

□

10.4 高阶近似

本节我们考虑稳定子流形的整体Taylor 展开. 在后面微分的计算中, 我们必须计算Φ_k, 为此我们需要下面的命题.

命题 10.4.1

$$W_{[n^s,n]} = \prod_{i=0}^{s-1} \left(I_{n^i} \otimes W_{[n,n]} \otimes I_{n^{s-i-1}}\right). \tag{10.4.1}$$

证明 利用命题1.5.11, 我们有

$$W_{[n^s,n]} = \left(W_{[n^{s-1},n]} \otimes I_n\right)\left(I_{n^{s-1}} \otimes W_{[n,n]}\right).$$

再次利用命题1.5.11 的公式分解第一个因子, 如此继续下去, 就得到(10.4.1). 注意, 这里我们约定$I_{n^0} = 1$, $\Phi_0 = I_n$. □

利用(10.4.1), 可以很容易计算Φ_k. 我们给出下面的例子说明.

例 10.4.2 设$n = 2$, 则

$$\Phi_0 = I_n,$$

$$\Phi_1 = W_{[n,n]} + I \otimes W_{[1,n]} = W_{[n]} + I_{n^2} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

$$\begin{aligned} \Phi_2 &= W_{[n^2,n]} + I_n \otimes W_{[n,n]} + I_{n^2} \otimes W_{[1,n]} \\ &= \left(W_{[n]} \otimes I_n\right)\left(I_n \otimes W_{[n]}\right) + I_n \otimes W_{[n]} + I_{n^3} \\ &= \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix}, \quad \cdots. \end{aligned}$$

□

接下来, 我们将从(10.2.4)~(10.2.6) 中解出H_k. 问题是x^k 是k 次齐次多项式的一组冗余基底, 因此从(10.2.4)~(10.2.6) 我们不能得到唯一解. 为了克服这个困难, 我们考虑k 次齐次多项式的自然基底. 回忆第5 章, 设$S \in \mathbb{Z}_+^n$, 自然基底可以定义为

$$B_n^k = \{x^S \mid S \in \mathbb{Z}_+^n,\ |S| = k\}.$$

现在我们按照字母序排列B_n^k 中的元素, 即, 对于$S^1 = (s_1^1, \cdots, s_n^1)$ 和$S^2 = (s_1^2, \cdots, s_n^2)$ 我们记次序为$x^{S^1} \prec x^{S^2}$ 如果存在一个t, $1 \le t \le n-1$, 使得

$$s_1^1 = s_1^2,\ \cdots,\ s_t^1 = s_t^2,\ s_{t+1}^1 > s_{t+2}^2.$$

这样我们将B_n^k 中的元素排成一列, 并且记它为$x_{(k)}$.

例 10.4.3 设$n=3$, $k=2$, 则

$$x^2=(x_1^2,x_1x_2,x_1x_3,x_2x_1,x_2^2,x_2x_3,x_3x_1,x_3x_2,x_3^2)^{\mathrm{T}},$$

且

$$x_{(2)}=(x_1^2,x_1x_2,x_1x_3,x_2^2,x_2x_3,x_3^2)^{\mathrm{T}}.$$

□

在第 5 章曾证明, B_n^k 的大小是

$$|B_n^k|:=d=\frac{(n+k-1)!}{k!(n-1)!},\quad k\geq 0,\quad n\geq 1. \tag{10.4.2}$$

同时, 定义两个矩阵$T_N(n,k)\in M_{n^k\times d}$ 和$T_B(n,k)\in M_{d\times n^k}$, 它们可以将一个基底转化为另一个, 即

$$x^k=T_N(n,k)x_{(k)},\quad x_{(k)}=T_B(n,k)x^k,$$

且

$$T_B(n,k)T_N(n,k)=I_d.$$

我们用一个例子来说明.

例 10.4.4 设$n=2$, $k=3$. 我们构造$T_B(n,k)$ 为

$$T_B(2,3)=\begin{array}{c}\begin{array}{cccccccc}(111)&(112)&(121)&(122)&(211)&(212)&(221)&(222)\end{array}\\ \begin{bmatrix}1&0&0&0&0&0&0&0\\0&1/3&1/3&0&1/3&0&0&0\\0&0&0&1/3&0&1/3&1/3&0\\0&0&0&0&0&0&0&1\end{bmatrix}\end{array}\begin{array}{c}(111)\\(112)\\(122)\\(222)\end{array}. \tag{10.4.3}$$

同理, 我们有

$$T_N(2,3)=\begin{array}{c}\begin{array}{cccc}(111)&(112)&(122)&(222)\end{array}\\ \begin{bmatrix}1&0&0&0\\0&1&0&0\\0&1&0&0\\0&0&1&0\\0&1&0&0\\0&0&1&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\end{array}\begin{array}{c}(111)\\(112)\\(121)\\(122)\\(211)\\(212)\\(221)\\(222)\end{array}. \tag{10.4.4}$$

□

回忆(10.3.1), 除了解H_k, 我们可以尝试去解G_k, 其中

$$H_k x^k = G_k x_{(k)}.$$

称F_k 是一个对称系数集合, 如果x^k 的两个元素相同, 则它们的系数也相同. 我们用下面的例子来说明.

例 10.4.5　设$n = 3$, $k = 2$, 则x^2 如例10.4.3. 对于一个给定的k 次齐次多项式$p(x) = x_1^2 + 2x_1x_2 - 3x_1x_3 + x_2^2 - x_3^2$, 我们可以将它表示成

$$p(x) = H_1 x^2 = (1,\ 2,\ -3,\ 0,\ 1,\ 0,\ 0,\ 0,\ -1)x^2.$$

另外, 我们也可以将它表示成

$$p(x) = H_2 x^2 = \left(1,\ 1,\ -\frac{3}{2},\ 1,\ 1,\ 0,\ -\frac{3}{2},\ 0,\ -1\right) x^2.$$

于是H_1 不是对称的, 而H_2 是对称的. □

直接计算可以得到

命题 10.4.6　*对称系数集合H_k 是唯一的, 而且*

$$H_k = G_k T_B(n,k), \quad G_k = H_k T_N(n,k). \tag{10.4.5}$$

现在我们考虑稳定子流形的方程$h(x)$ 的高阶项. 记

$$f(x) = F_1 x + F_2 x^2 + \cdots;$$

$$h(x) = H_1 x + H_2 x^2 + \cdots.$$

注意到$F_1 = J_f(0) = J$, $H_1 = \eta^{\mathrm{T}}$, 且H_2 由(10.3.4) 唯一决定.

命题 10.4.7　*$h(x)$ 的系数H_k, $k \geq 2$ 满足下面的方程*

$$\left[\sum_{i=1}^{k} H_i \Phi_{i-1}(I_{n^{i-1}} \otimes F_{k-i+1}) - \mu H_k\right] x^k = 0, \quad k \geq 2. \tag{10.4.6}$$

证明　注意到$h(x) = 0$ 是向量场$f(x)$ 不变的, 即李导数

$$L_f h(x) = 0. \tag{10.4.7}$$

利用(5.5.4), 我们有

$$Dh(x) = H_1 + H_2\Phi_1 x + H_3\Phi_2 x^2 + \cdots = H_1 + 2x^{\mathrm{T}}\Psi + H_3\Phi_2 x^2 + \cdots.$$

直接计算有

$$\begin{aligned} L_f h(x) \quad &= \quad \mu\eta^{\mathrm{T}} x + [H_2\Phi_1(I_n \otimes F_1) + H_1 F_2]\, x^2 + \cdots \\ &\quad + \left[\sum_{i=1}^{k} H_i\Phi_{i-1}(I_{n^{i-1}} \otimes F_{k+1-i})\right] x^k + \cdots. \end{aligned}$$

注意到稳定子流形的方程满足

$$\begin{cases} h(x) = 0, \\ L_f h(x) = 0. \end{cases} \tag{10.4.8}$$

将(10.4.8) 的第二个方程减去μ 和(10.4.8) 的第一个方程的乘积, 对k 进行归纳我们有

$$\left[\sum_{i=1}^{k} H_i\Phi_{i-1}(I_{n^{i-1}} \otimes F_{k-i+1}) - \mu H_k\right] x^k + O(\|x\|^{k+1}) = 0, \quad k \geq 2,$$

于是命题得证. □

观察(10.4.6), 根据命题10.4.6, 它可以表示成

$$\begin{aligned} &G_k\left[\mu I_d - \mathrm{T}_B(n,k)\Phi_{k-1}(I_{n^{k-1}} \otimes F_1)\mathrm{T}_N(n,k)\right] x_{(k)} \\ &\equiv \left[\sum_{i=1}^{k-1} G_i \mathrm{T}_B(n,i)\Phi_{i-1}(I_{n^{i-1}} \otimes F_{k-i+1})\right] \mathrm{T}_N(n,k) x_{(k)}, \quad k \geq 3. \end{aligned} \tag{10.4.9}$$

下面的定理是上述论述的总结, 适用于一般情况.

定理 10.4.8 设矩阵

$$C_k := \mu I_d - \mathrm{T}_B(n,k)\Phi_{k-1}(I_{n^{k-1}} \otimes F_1)\mathrm{T}_N(n,k), \quad k \geq 3 \tag{10.4.10}$$

非奇异, 则

$$G_k = \left[\sum_{i=1}^{k-1} G_i \mathrm{T}_B(n,i)\Phi_{i-1}(I_{n^{i-1}} \otimes F_{k-i+1})\right] \mathrm{T}_N(n,k) C_k^{-1}. \tag{10.4.11}$$

注 实际上, H_2 也可以用这种方法解出. (10.3.2) 和(10.4.11) 得到的结果相同. 实际上, 当H_2 由(10.3.2) 解出时, 由于使用了对称的二次型, 系数的对称性已经自动考虑进去了.

显然(10.4.11) 是否有效的关键是C_i 是否非奇异. 不幸的是, 这不是确定的. 我们用下面的例子说明这个问题, 并讨论进一步解决这个问题的方法.

例 10.4.9 考虑下面的系统

$$\begin{cases}\dot{x}_1 = -cx_1, \quad c > 0,\\ \dot{x}_2 = x_2 - 2x_1^2 + x_1^3,\end{cases} \tag{10.4.12}$$

其中$c > 0$ 是一个参数.

我们计算稳定子流形. 易算出$\mu = 1$, $\eta = (0\ 1)^{\mathrm{T}}$,

$$J = \begin{bmatrix} -c & 0 \\ 0 & 1 \end{bmatrix},$$

且

$$\mathrm{Hess}(f_1(0)) = 0, \quad \mathrm{Hess}(f_2(0)) = \begin{bmatrix} -4 & 0 \\ 0 & 0 \end{bmatrix}.$$

因此我们可以用(10.3.10) 算出

$$h(x) = (0\ 1)x + x^{\mathrm{T}} \begin{bmatrix} -\frac{2}{2c+1} & 0 \\ 0 & 0 \end{bmatrix} x + O(\|x\|^3). \tag{10.4.13}$$

利用(10.4.2), (10.4.3) 和例10.4.2 中的Φ_2, 我们可以算出C_3 为

$$C_3 = \begin{bmatrix} 3c+1 & 0 & 0 & 0 \\ 0 & 2c & 0 & 0 \\ 0 & 0 & c-1 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}. \tag{10.4.14}$$

假设$c \neq 1$, 则C_3 是可逆的. 由上面我们有

$$H_1 = (0,\ 1), \quad H_2 = \left(-\frac{2}{2c+1},\ 0,\ 0,\ 0\right),$$

$$F_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & 0 \end{bmatrix}, \quad F_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

将它们代入(10.4.10) 得到

$$G_3 = \left(\frac{1}{3c+1},\ 0,\ 0,\ 0\right).$$

于是得到

$$h(x)=x_2-\frac{2}{2c+1}x_1^2+\frac{1}{3c+1}x_1^3+O(\|x\|^4).$$

实际上, 容易验证

$$h(x)=x_2-\frac{2}{2c+1}x_1^2+\frac{1}{3c+1}x_1^3=0,$$

我们还有

$$L_f h(x)=h(x)=0.$$

因此

$$W^s(0)=\left\{x\in\mathbb{R}^2\,\middle|\,x_2-\frac{2}{2c+1}x_1^2+\frac{1}{3c+1}x_1^3=0\right\}.$$

□

根据定理10.4.8 和例10.4.9, 我们给出下面的算法:

第1步: 如果$C_3,\cdots,C_{k-1}$ 非奇异, 我们可以继续寻找H_k 来更好地近似$h(x)$ 直到满足精度的要求.

第2步: 如果C_k 奇异, 我们可以寻找

$$\begin{aligned}&G_k\left[\mu I_d-\mathrm{T}_B(n,k)\Phi_{k-1}(I_{n^{k-1}}\otimes F_1)\mathrm{T}_N(n,k)\right]\\&=\left[\sum_{i=1}^{k-1}G_i\mathrm{T}_B(n,i)\Phi_{i-1}(I_{n^{i-1}}\otimes F_{k-i+1})\right]\mathrm{T}_N(n,k)\end{aligned}\tag{10.4.15}$$

的最小二乘解G_k, 并且固定k 作为近似$h(x)$ 的阶数.

第3步(可能的进一步改进): 如果得到的最小二乘解是(10.4.15) 的一个实数解, 解由下面的方程构成的系统:

$$\begin{cases}G_k\left[\mu I_d-\mathrm{T}_B(n,k)\Phi_{k-1}(I_{n^{k-1}}\otimes F_1)\mathrm{T}_N(n,k)\right]\\ \quad=\left[\displaystyle\sum_{i=1}^{k-1}G_i\mathrm{T}_B(n,i)\Phi_{i-1}(I_{n^{i-1}}\otimes F_{k-i+1})\right]\mathrm{T}_N(n,k),\\ 0=\left[\displaystyle\sum_{i=1}^{k}G_i\mathrm{T}_B(n,i)\Phi_{i-1}(I_{n^{i-1}}\otimes F_{k-i+1})\right]\mathrm{T}_N(n,k+1).\end{cases}\tag{10.4.16}$$

实际上, 通过同时考虑k 阶项和$k+1$ 阶项得到(10.4.16).

回顾例10.4.9. 当$c=1$ 时, 最小二乘解是

$$G_3=\left(\frac{1}{3c+1},\ 0,\ t,\ 0\right),$$

其中t 是一个任意参数. 可以验证G_3 是(10.4.15) 的一个实数解. 因此我们可以尝试解(10.4.16). 通过仔细的计算可以证明(10.4.16) 的解是$G_3=\left(\frac{1}{3c+1},\ 0,\ 0,\ 0\right)$. 容易验证这个$G_3$ 是一个实数解.

下面我们用一个更一般的例子来说明这个算法.

例 10.4.10 考虑下面的系统

$$\begin{cases}\dot{x}_1=x_2,\\ \dot{x}_2=-x_1-2x_2,\\ \dot{x}_3=2x_3-x_2(e^{x_1}-1).\end{cases}\tag{10.4.17}$$

简单地计算有$\mu=2$, $\eta=(0\ 0\ 1)^{\mathrm{T}}$,

$$J=\begin{bmatrix}0&1&0\\-1&-2&0\\0&0&2\end{bmatrix},$$

$$A=\frac{\mu}{2}I_3-J^{\mathrm{T}}=\begin{bmatrix}1&1&0\\-1&3&0\\0&0&-1\end{bmatrix},$$

$$\mathrm{Hess}(f_1(0))=\mathrm{Hess}(f_2(0))=0,$$

$$\mathrm{Hess}(f_3(0))=\begin{bmatrix}0&-1&0\\-1&0&0\\0&0&0\end{bmatrix}.$$

利用公式(10.3.9), 我们有

$$\begin{aligned}h(x)&\approx\eta^{\mathrm{T}}x+x^{\mathrm{T}}(\tfrac{1}{2}\Psi)x\\&=(0\ 0\ 1)x+x^{\mathrm{T}}\begin{bmatrix}0.09375&-0.09375&0\\-0.09375&-0.03125&0\\0&0&0\end{bmatrix}x\\&=x_3+0.09375x_1^2-0.1875x_1x_2-0.03125x_2^2.\end{aligned}\tag{10.4.18}$$

现在我们继续计算三次项. 首先我们检查C_3. 利用(10.4.10), 我们有

$$C_3=\begin{bmatrix}2&-3&0&0&0&0&0&0&0&0\\1&4&0&-2&0&0&0&0&0&0\\0&0&0&0&-2&0&0&0&0&0\\0&2&0&6&0&0&-1&0&0&0\\0&0&1&0&2&0&0&-1&0&0\\0&0&0&0&0&-2&0&0&-1&0\\0&0&0&3&0&0&8&0&0&0\\0&0&0&0&2&0&0&4&0&0\\0&0&0&0&0&1&0&0&0&0\\0&0&0&0&0&0&0&0&0&-4\end{bmatrix}.$$

它可以通过计算机算出, 并且通过计算机容易验证它是可逆的. 从$h(x)$ 的二次部分我们还有

$$H_1=\eta^{\mathrm{T}}=(0,\ 0,\ 1),$$

$$H_2=(0.09375,\ -0.09375,\ 0,\ -0.09375,\ -0.03125,\ 0,\ 0,\ 0,\ 0)$$

$F_2\in M_{3\times 9}$ 除$F_2(3,2)$ 和$F_2(3,4)$ 外其他元素均为零, 非零元素为

$$F_2(3,2)=F_2(3,4)=-\frac{1}{2},$$

$F_3\in M_{3\times 29}$ 只有3 个非零元素: $F_3(3,2)$, $F_3(3,4)$ 和$F_3(3,10)$, 且

$$F_3(3,2)=F_3(3,4)=F_3(3,10)=-\frac{1}{6}.$$

将它们全部代入(10.4.11) 得到

$$G_3=(0.0408,\ -0.0816,\ 0,\ -0.0256,\ 0,\ 0,\ -0.0032,\ 0,\ 0,\ 0).$$

因此稳定子流形的函数的近似到3 次方的项为

$$\begin{aligned}h(x)\approx\ & x_3+0.09375x_1^2-0.1875x_1x_2-0.03125x_2^2+0.0408x_1^3\\&-0.0816x_1^2x_2-0.0256x_1x_2^2-0.0032x_2^3.\end{aligned}\tag{10.4.19}$$

我们继续下去就可以得到$h(x)$ 的更高阶项.

实际上, 对于这个特殊的系统, 可以通过坐标变换直接算出它的稳定子流形. 因此可以验证上面得到的结果是正确的. □

在结束本节前, 作为例子, 我们将这个算法应用于一个实际的物理系统.

例 10.4.11 [70] 一个单摆系统可以描述为

$$\begin{cases}\dot{x}_1 = x_2; \\ \dot{x}_2 = -\sin x_1 - 0.5x_2;\end{cases} \tag{10.4.20}$$

图10.4.1~图10.4.3 分别给出了它的稳定子流形的二阶、三阶和五阶近似. □

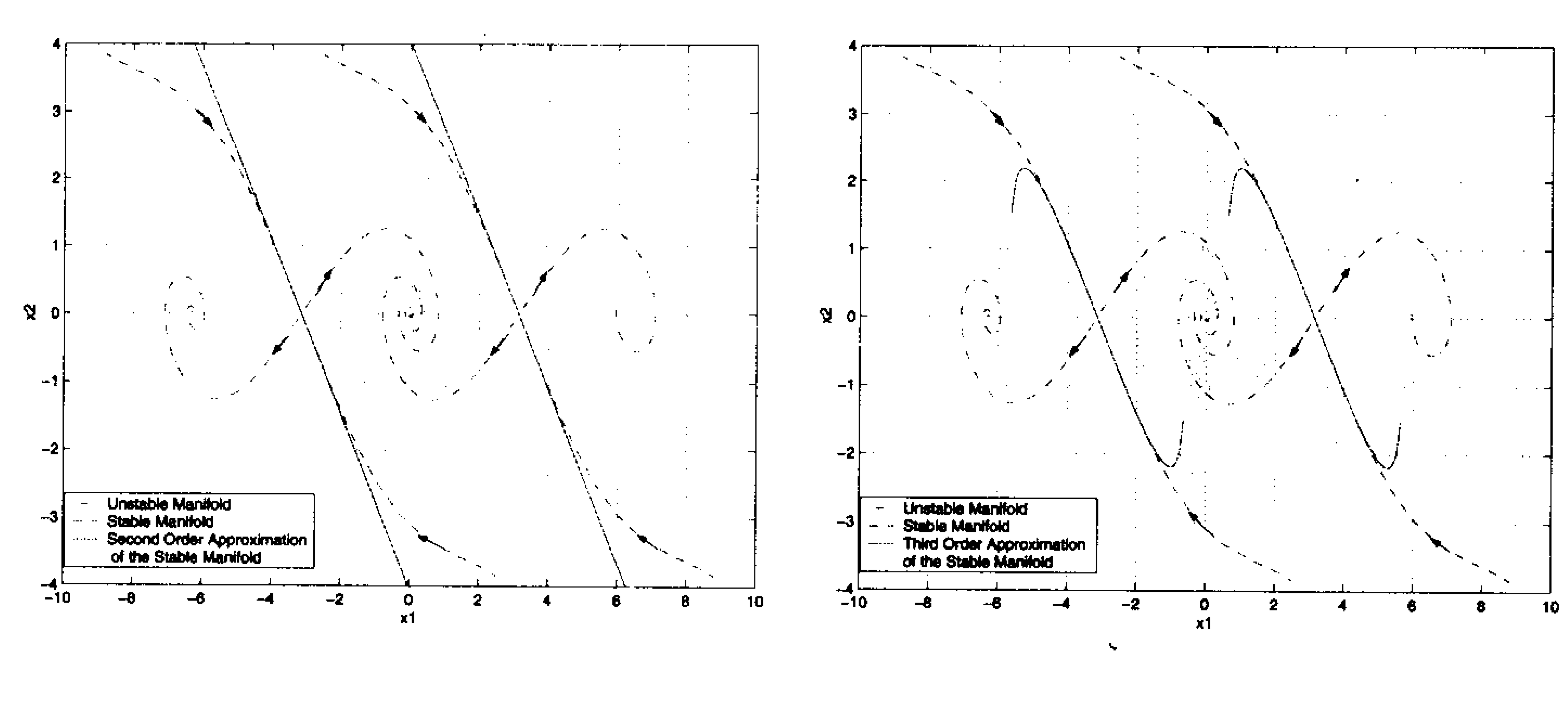

图 10.4.1 二阶近似

图 10.4.2 三阶近似

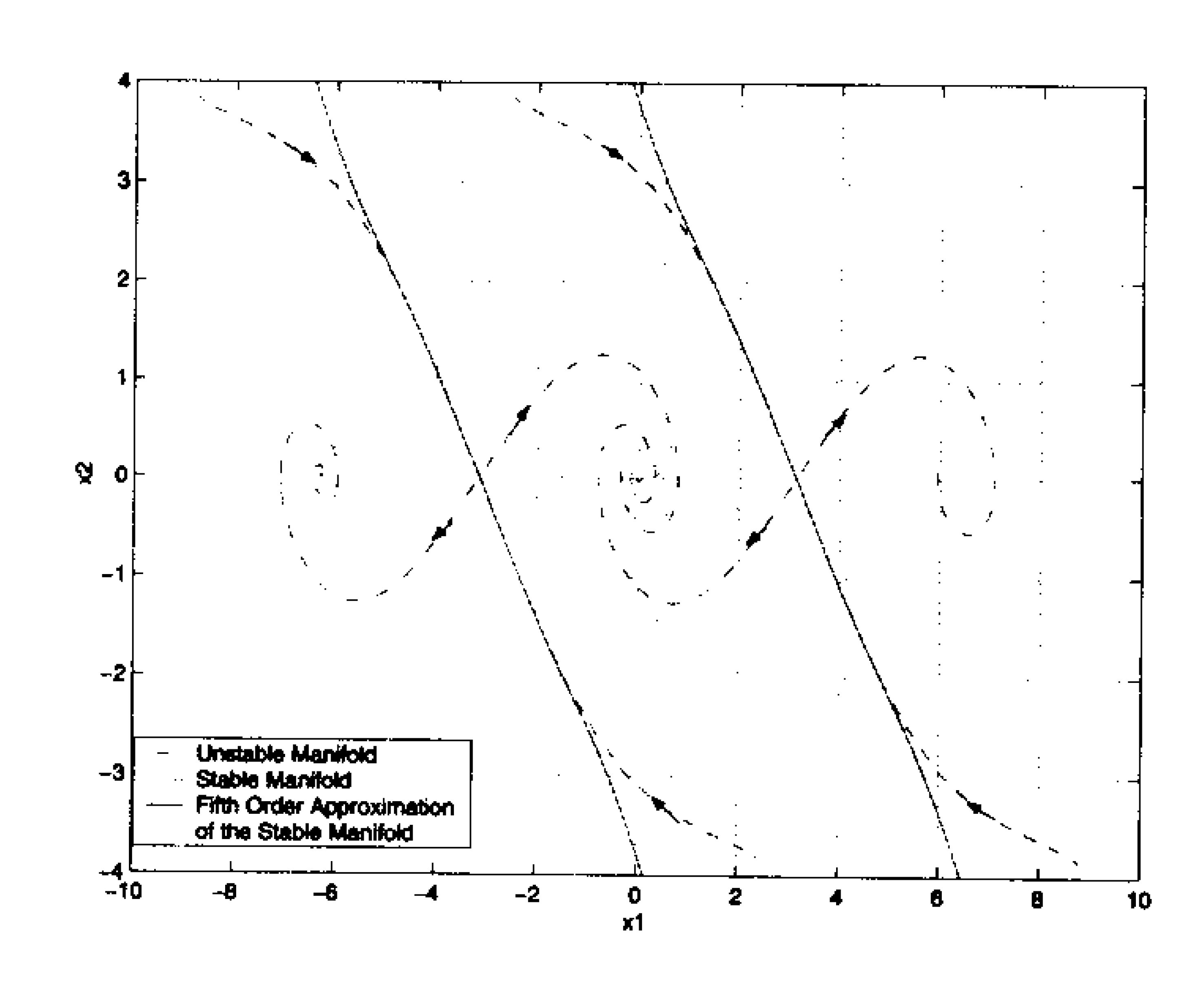

图 10.4.3 五阶近似

10.5 微分–代数系统

本节考虑微分–代数系统下稳定域的边界的计算. 这类系统在工程问题中大

量存在, 电网系统就是一个典型例子. 考虑如下系统:

$$\begin{cases}\dot{x} = f(x,y), & x \in \mathbb{R}^n,\ y \in \mathbb{R}^m, \\ \Phi(x,y) = 0, & \Phi(x,y) \in \mathbb{R}^m,\end{cases} \tag{10.5.1}$$

这里$f(0,0)=0$, $\Phi(0,0)=0$. 并且,我们假定动力系统是唯一确定的, 因此要求

$$\operatorname{rank}\left(\frac{\partial \Phi}{\partial y}(0,0)\right) = m. \tag{10.5.2}$$

与前几节一样, 我们假定$(0,0)$ 是1 型不稳定平衡点. 我们将利用前面的结果给出其稳定子流形的方程. 为方便计, 我们只考虑系统(10.5.1) 在该点的二次近似表达式, 高次近似式可类似地导出.

注意到, 根据(10.5.2) 由隐函数存在定理可知, y 可由(10.5.1) 的第二式局部解出, 于是可得$y=y(x)$. 将其代入(10.5.1), 即可得到一个形同(10.1.1) 的标准微分方程

$$\dot{x} = f(x, y(x)). \tag{10.5.3}$$

虽然这种表示是局部的, 但Taylor 展式表示其实只需要单点信息, 因此, 局部信息已经够了. 唯一的困难在于解出$y=y(x)$, 一般情况下这是做不到的. 回忆(10.3.2), 我们真正需要求的是

$$\begin{aligned} &J := \tfrac{\partial f(x,y(x))}{\partial x}, \\ &H_i := Hess(f_i(0,0)), \quad i=1,\cdots,n. \end{aligned} \tag{10.5.4}$$

如果能算出J 和H_i, 则可直接使用公式(10.3.2) 了.

因为

$$\frac{\partial \Phi}{\partial x} + \frac{\partial \Phi}{\partial y}\frac{\partial y}{\partial x} = 0,$$

则

$$\frac{\partial y}{\partial x} = -\left(\frac{\partial \Phi}{\partial y}\right)^{-1}\frac{\partial \Phi}{\partial x}. \tag{10.5.5}$$

根据复合求导公式, 有

$$J = \frac{\partial f}{\partial x}(0,0) - \frac{\partial f}{\partial y}(0,0)\left(\frac{\partial \Phi}{\partial y}(0,0)\right)^{-1}\frac{\partial \Phi}{\partial x}(0,0). \tag{10.5.6}$$

由第5 章(5.2.10), 设$A(x)$, $B(x)$ 分别为$p\times q$ 及$q\times r$ 函数矩阵, 则

$$DA(x)B(x)=DA(x)B(x)+A(x)DB(x). \tag{10.5.7}$$

同样, 据复合求导公式, 有

$$DA(x,y(x))=D_xA(x,y)+D_yA(x,y)\left(I_n\otimes\frac{\partial y}{\partial x}\right). \tag{10.5.8}$$

这里, 我们用D_x 表示仅对x 作微分.

下面计算H_i: 首先, f_i 的梯度可表示为

$$\begin{aligned}\nabla f_i(x,y(x)) &= \nabla_x f_i(x,y)+\left(d_y f_i(x,y)\tfrac{\partial y}{\partial x}\right)^{\mathrm T}\\ &= \nabla_x f_i(x,y)-\left(\tfrac{\partial\Phi}{\partial x}\right)^{\mathrm T}\left(\tfrac{\partial\Phi}{\partial y}\right)^{-\mathrm T}\nabla_y f_i(x,y).\end{aligned} \tag{10.5.9}$$

考虑$y=y(x)$ 为x 的一个函数, 同时, $\nabla_x f_i(x,y)$ 和$\nabla_y f_i(x,y)$ 分别为相对于x 和y 的梯度.

于是, 依定义我们有

$$H_i=D(\nabla f_i)|_{(0,0)},\quad i=1,\cdots,n. \tag{10.5.10}$$

应用(10.5.8) 到(10.5.9) 的第一式, 我们有

$$D(\nabla_x f_i)=\frac{\partial^2 f_i}{\partial x\partial x}+\frac{\partial^2 f_i}{\partial x\partial y}\left(\frac{\partial y}{\partial x}\right). \tag{10.5.11}$$

类似地, 我们有

$$D(\nabla_y f_i)=\frac{\partial^2 f_i}{\partial y\partial x}+\frac{\partial^2 f_i}{\partial y\partial y}\left(\frac{\partial y}{\partial x}\right). \tag{10.5.12}$$

注意到对任一函数$\xi(x,y)$, 我们用$\frac{\partial^2\xi(x,y)}{\partial x\partial y}$ 表示一个$n\times m$ 函数矩阵, 其(i,j) 元素为$\frac{\partial^2\xi}{\partial x_i\partial y_j}$. 因此, 一般地说$\frac{\partial^2\xi(x,y)}{\partial x\partial y}\neq\frac{\partial^2\xi(x,y)}{\partial y\partial x}$.

将(10.5.8) 应用于(10.5.9) 的第二项, 可得

$$\begin{aligned}&D\left[d_y f_i(x,y)\tfrac{\partial y}{\partial x}\right]^{\mathrm T}=D\left\{\left(\tfrac{\partial y}{\partial x}\right)^{\mathrm T}(\nabla_y f_i(x,y))\right\}\\ &=D\left(\tfrac{\partial y}{\partial x}\right)^{\mathrm T}(\nabla_y f_i(x,y)\otimes I_n)+\left(\tfrac{\partial y}{\partial x}\right)^{\mathrm T}D(\nabla_y f_i(x,y)).\end{aligned} \tag{10.5.13}$$

我们逐项计算(10.5.13). 利用(10.5.5), 我们有

$$\left(\frac{\partial y}{\partial x}\right)^{\mathrm T}\left(\frac{\partial\Phi}{\partial y}\right)^{\mathrm T}+\left(\frac{\partial\Phi}{\partial x}\right)^{\mathrm T}=0,$$

对上式两边微分, 并利用(10.5.7), 可得

$$D\left(\frac{\partial y}{\partial x}\right)^{\mathrm{T}}\left[\left(\frac{\partial \Phi}{\partial y}\right)^{\mathrm{T}} \otimes I_n\right]+\left(\frac{\partial y}{\partial x}\right)^{\mathrm{T}} D\left(\frac{\partial \Phi}{\partial y}\right)^{\mathrm{T}}+D\left(\frac{\partial \Phi}{\partial x}\right)^{\mathrm{T}}=0. \qquad (10.5.14)$$

直接计算各项如下:

$$\begin{aligned} X &:= D\left(\tfrac{\partial \Phi}{\partial x}\right)^{\mathrm{T}}\Big|_{(0,0)} \\ &= \left[\left(\tfrac{\partial^2 \Phi_1}{\partial x \partial x}+\tfrac{\partial^2 \Phi_1}{\partial x \partial y} \cdot \tfrac{\partial y}{\partial x}\right), \cdots,\left(\tfrac{\partial^2 \Phi_m}{\partial x \partial x}+\tfrac{\partial^2 \Phi_m}{\partial x \partial y} \cdot \tfrac{\partial y}{\partial x}\right)\right]\Big|_{(0,0)}. \end{aligned} \qquad (10.5.15)$$

$$\begin{aligned} Y &:= D\left(\tfrac{\partial \Phi}{\partial y}\right)^{\mathrm{T}}\Big|_{(0,0)} \\ &= \left[\left(\tfrac{\partial^2 \Phi_1}{\partial y \partial x}+\tfrac{\partial^2 \Phi_1}{\partial y \partial y} \cdot \tfrac{\partial y}{\partial x}\right), \cdots,\left(\tfrac{\partial^2 \Phi_m}{\partial y \partial x}+\tfrac{\partial^2 \Phi_m}{\partial y \partial y} \cdot \tfrac{\partial y}{\partial x}\right)\right]\Big|_{(0,0)}. \end{aligned} \qquad (10.5.16)$$

将它们代入(10.5.14) 即可得到

$$D\left(\frac{\partial y}{\partial x}\right)^{\mathrm{T}}\Big|_{(0,0)}=-\left[\left(\frac{\partial y}{\partial x}(0,0)\right)^{\mathrm{T}} Y+X\right]\left[\left(\frac{\partial \Phi}{\partial y}(0,0)\right)^{-\mathrm{T}} \otimes I_n\right]. \qquad (10.5.17)$$

最后,将(10.5.11), (10.5.12) 及(10.5.17) 代入(10.5.10), 就可得到H_i 的如下表达式

$$\begin{aligned} H_i = & \tfrac{\partial^2 f_i}{\partial x \partial x}+\tfrac{\partial^2 f_i}{\partial x \partial y}\left(\tfrac{\partial y}{\partial x}\right)-\left[\left(\tfrac{\partial y}{\partial x}\right)^{\mathrm{T}} Y+X\right]\left[\left(\tfrac{\partial \Phi}{\partial y}\right)^{-\mathrm{T}} \otimes I_n\right]\left(\nabla_y f_i \otimes I_n\right) \\ & +\left(\tfrac{\partial y}{\partial x}\right)^{\mathrm{T}}\left[\tfrac{\partial^2 f_i}{\partial y \partial x}+\tfrac{\partial^2 f_i}{\partial y \partial y}\left(\tfrac{\partial y}{\partial x}\right)\right], \end{aligned} \qquad (10.5.18)$$

这里X, Y 以及$\frac{\partial y}{\partial x}$ 的具体表达式可分别见(10.5.15), (10.5.16) 以及(10.5.5).

10.6 注释与参考

关于电力系统稳定域方面的文章甚多, 与本章思路相关的有文献[37], [35].

关于能量平衡法, 可见文献[38].

电力系统的吸引域几何结构的理论基础是由文献[100], [122], [35], [36] 中的工作建立的. 它们是本章讨论的出发点.

关于1 型平衡点的近似算法很多; 例如, 在文献[101] 中用动力系统的Taylor展开用来得到一个二次近似, 而在文献[91] 中介绍了通过向量场的正则型得到一个稳定域边界的二次近似.

本章主要内容前四节来自文献[25], 第5节来自文献[29]. 同时,在文献[25]和[29] 中各有若干在电力系统中应用的例子.

用半张量积方法计算稳定域边界的工作的一个新进展可见文献[116], [117], [118], [119], [120].

第 11 章　Morgan 问题

本章主要考虑半张量积在输入输出解耦问题即Morgan 问题中的应用. 首先, 回顾控制系统的Morgan 问题, 然后对Morgan 问题进行恰当的描述, 为问题的代数化作准备, 最后将Morgan 问题转化为一个代数问题, 并且给出了一个算法.

11.1　输入输出解耦

考虑线性系统

$$\begin{cases} \dot{x} = Ax + \sum_{i=1}^{m} b_i u_i := Ax + Bu, \quad x \in \mathbb{R}^n,\ u \in \mathbb{R}^m, \\ y = Cx, \quad y \in \mathbb{R}^p. \end{cases} \tag{11.1.1}$$

其中$m \geq p$. 一般我们假设$\operatorname{rank}(B) = m$, $\operatorname{rank}(C) = p$. 控制理论中的一个重要问题是输入输出解耦问题, 即每个输出仅由一个输入所完全控制, 且不同的输出由不同的输入所控制. 更确切些, 问题可以叙述为

定义 11.1.1　*考虑系统(11.1.1). 输入输出解耦问题(Morgan 问题) 可解, 如果存在反馈*

$$u = kx + Hv, \quad v \in \mathbb{R}^p, \tag{11.1.2}$$

使得闭环系统的传递矩阵是对角阵且非奇异.

为了讨论的方便, 先给出下面的一些定义: 对于每个输出$y_j = c_j x$, 定义相对阶ρ_j 为

$$\rho_j = \min\{i \,|\, c_j A^{i-1} B \neq 0\}, \quad j = 1, \cdots, p. \tag{11.1.3}$$

利用相对阶向量$(\rho_1, \cdots, \rho_p)$, 定义$p \times m$ 解耦矩阵为

$$D = \begin{bmatrix} c_1 A^{\rho_1 - 1} B \\ c_2 A^{\rho_2 - 1} B \\ \vdots \\ c_p A^{\rho_p - 1} B \end{bmatrix}. \tag{11.1.4}$$

对于$m = p$ 的情况, 我们有下面的经典结果:

定理 11.1.2 [47]　*当$m = p$ 时, Morgan 问题可解, 当且仅当解耦矩阵D 非奇异.*

相应于线性系统, 我们来看看非线性系统的Morgan 问题. 考虑仿射非线性系统

$$\begin{cases} \dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i := f(x) + G(x)u, \quad x \in \mathbb{R}^n, \ u \in \mathbb{R}^m, \\ y_j = h_j(x), \quad j = 1, \cdots, p. \end{cases} \tag{11.1.5}$$

这里$f(x)$, $g_i(x)$ $i = 1, \cdots, m$ 是光滑向量场, $h_j(x)$, $j = 1, \cdots, p$ 是光滑函数. 对于系统(11.1.5), 定义相对阶向量$\rho = (\rho_1, \cdots, \rho_p)^{\mathrm{T}}$ 为

$$\begin{aligned} &L_{g_i} L_f^k h_j(x) = 0, \quad x \in U, \\ &\qquad i = 1, \cdots, m; \ k = 0, 1, \cdots, \rho_j - 2. \\ &L_{g_i} L_f^{\rho_j - 1} h_j(x_0) \neq 0, \quad \exists i \in \{1, 2, \cdots, m\}. \end{aligned}$$

其中U 是x_0 的一个邻域, x_0 是我们要考虑的点. 为了简单起见, 可令$x_0 = 0$. 假设相对阶向量有定义, 且定义解耦矩阵为

$$D(x) = \begin{bmatrix} L_{g_1} L_f^{\rho_1 - 1} h_1(x) & \cdots & L_{g_m} L_f^{\rho_1 - 1} h_1(x) \\ L_{g_1} L_f^{\rho_2 - 1} h_2(x) & \cdots & L_{g_m} L_f^{\rho_2 - 1} h_2(x) \\ \vdots & & \vdots \\ L_{g_1} L_f^{\rho_p - 1} h_p(x) & \cdots & L_{g_m} L_f^{\rho_p - 1} h_p(x) \end{bmatrix}. \tag{11.1.6}$$

相应于线性情形, 我们有下面的结论:

定理 11.1.3 [49] 当$m = p$ 时, 非线性系统(11.1.5) 的Morgan 问题在x_0 处局部可解, 当且仅当解耦矩阵$D(x_0)$ 非奇异.

11.2 简化的等价形式

由上一节可知, 真正需要研究的问题是$m > p$ 的情形. 由定理11.1.2, 显然有下面的引理.

引理 11.2.1 Morgan 问题可解, 当且仅当存在$K \in M_{m\times n}$, $H \in M_{m\times p}$, $1 \leq \rho_i \leq n$, $i = 1, \cdots, p$ 使得

$$c_i(A + BK)^{t_i} BH = 0, \quad t_i = 0, \cdots, \rho_i - 2; \quad i = 1, \cdots, p. \tag{11.2.1}$$

且

$$D=\begin{bmatrix} c_1(A+BK)^{\rho_1-1}BH \\ \vdots \\ c_p(A+BK)^{\rho_p-1}BH \end{bmatrix} \tag{11.2.2}$$

是非奇异的.

这里有两个可选择的反馈矩阵K 和H, 本节的目标是给出一个等价条件, 它将未知矩阵个数减为一个.

首先, 注意到$\rho_i = 1$ 时(11.2.1) 不存在, 记

$$\Lambda = \{i \mid \rho_i = 2\}, \quad C_\Lambda = col\{c_i \mid i \in \Lambda\}.$$

于是, 根据相对阶的定义, 当$1 \le \rho_i \le 2$, $i = 1, \cdots, p$, (11.2.1) 变为

$$H \subset (C_\Lambda B)^\perp.$$

记

$$\Lambda^c = \{1, \cdots, p\} \backslash \Lambda.$$

则$i \in \Lambda^c$ 意味着$\rho_i = 1$.

因此我们有

推论 11.2.2 对于$1 \le \rho_i \le 2$, $i = 1, \cdots, p$, Morgan 问题可解, 当且仅当存在$K \in M_{m \times n}$, 使得

$$D=\begin{bmatrix} C_{\Lambda^c}B \\ c_\Lambda(A+BK)B \end{bmatrix}(C_\Lambda B)^\perp. \tag{11.2.3}$$

行满秩.

对于一般情况, 消去H 不那么容易, 需要更仔细的工作.

定义

$$W(K):=\begin{bmatrix} c_1B \\ c_1(A+BK)B \\ \vdots \\ c_1(A+BK)^{\rho_1-2}B \\ \vdots \\ c_p(A+BK)^{\rho_p-2}B \end{bmatrix}, \quad T(K):=\begin{bmatrix} c_1(A+BK)^{\rho_1-1}B \\ \vdots \\ c_p(A+BK)^{\rho_p-1}B \end{bmatrix}.$$

则(11.2.1) 变为

$$W(K)H=0, \tag{11.2.4}$$

(11.2.2) 变为

$$D=T(K)H. \tag{11.2.5}$$

由于$1\leq \rho_i\leq n$, $i=1,\cdots,p$, 对于固定的ρ_i, $i=1,\cdots,p$ 我们可以考虑Morgan 问题的可解性, 因为只需要验证有限(n^p) 种情形. 在本章的剩余部分, 如非特别说明, 我们都是在一组固定的ρ_i 下考虑Morgan 问题的可解性.

引理 11.2.3 Morgan 问题可解, 当且仅当存在$K\in M_{m\times n}$ 使得

(1)

$$\mathrm{Im}(T^{\mathrm{T}}(K))\cap \mathrm{Im}(W^{\mathrm{T}}(K))=\{0\}. \tag{11.2.6}$$

(2) $T(K)$ 行满秩.

证明 我们可以证明下列命题是等价的:

(i) 存在H 使得$T(K)H$ 非奇异, 且$W(K)H=0$;

(ii) $T(K)((W(K)^{\mathrm{T}})^{\perp})=\mathbb{R}^p$;

(iii) $[(T^{\mathrm{T}}(K))^{-1}(W(K)^{\mathrm{T}})]^{\perp}=\mathbb{R}^p$;

(iv) $(T^{\mathrm{T}}(K))^{-1}(W(K)^{\mathrm{T}})=0$;

(v) 条件(1) 和(2).

其中$T(K)$ 和$(T^{\mathrm{T}}(K))^{-1}$ 都看作函数映射[111].

(i)⇒(ii): 如果$\dim(T(k)(W(K)^{\mathrm{T}})^{\perp})<p$, 由于$\mathrm{Im}(H)\subset (W(K)^{\mathrm{T}})^{\perp}$, 那么$\mathrm{rank}(T(K)H)<p$, 这就产生了矛盾;

(ii)⇒(i): 选择p 个向量$h_i\in (W(K)^{\mathrm{T}})^{\perp}$, 使得$T(K)\mathrm{Im}(h_1,\cdots,h_p)=\mathbb{R}^p$. 于是令$H=(h_1,\cdots,h_p)$;

(ii)⇔(iii): 参见文献[111] 第23 页;

(iii)⇔(iv) 是显然的;

(iv)⇔(v): 容易验证(iv) 和(v) 都等价于如下命题: 如果$Y\in\mathbb{R}^p$ 且$T^{\mathrm{T}}(K)Y\in \mathrm{Im}((W(K))^{\mathrm{T}})$, 则$Y=0$. □

由上面的引理立即可以得到下面的定理.

定理 11.2.4 (对于固定的ρ_i) Morgan 问题可解, 当且仅当存在$K_0\in M_{m\times n}$, 使得

$$\mathrm{rank}\left(\begin{bmatrix}T(K_0)\\ W(K_0)\end{bmatrix}\right)=p+\mathrm{rank}(W(K_0)). \tag{11.2.7}$$

考虑$\rho_i \le 2, \forall i$ 的情况, 如果下述假定成立:

A1. $C_\Lambda B = 0$, 则无须考虑$W(K_0)$, 于是有

推论 11.2.5 设$1 \le \rho_i \le 2$, $i = 1, \cdots, p$, 并且A1 成立, 且存在$K_0 \in M_{m\times n}$ 使得$T(K_0)$ 行满秩, 则Morgan 问题可解.

定义 11.2.6 设$A(K)$ 是一个矩阵, 且它的元素$a_{ij}(K)$ 都是K 的多项式, 其中$K \in M_{m\times n}$.定义$A(K)$ 的本性秩$\text{rank}_e(A(K))$ 为

$$\text{rank}_e(A(K)) = \max_{K\in M_{m\times n}} \text{rank}(A(K)).$$

现在(在固定的ρ_i 下) 记

$$\text{rank}_e(T(K)) = t, \quad \text{rank}_e(W(K)) = s, \quad \text{rank}_e\left(\begin{bmatrix} T(K) \\ W(K) \end{bmatrix}\right) = q.$$

由于本性秩是很容易计算的, 下面的推论在一些情形下是很方便的.

推论 11.2.7 Morgan 问题可解, 如果$q = p + s$.

由于$T(K)$ 及$W(K)$ 均为K 的多项式矩阵, 本性秩在除K 的一个零测集外的其余K 处均能达到, 因此可以通过程序机械地计算本性秩, 这样很容易验证推论11.2.7.

11.3 可解性的代数表达

首先, 我们需要计算$T(K)$ 和$W(K)$. 记$Z = V_r(K) \in \mathbb{R}^{mn}$, 我们先将$T(K)$ 和$W(K)$ 表示成Z 的多项式形式. 两个矩阵的乘积可以用半张量积表示.

引理 11.3.1 给定矩阵$A \in M_{n\times m}$.

(1) 如果$x \in \mathbb{R}^n$ 是一个行向量, 则

$$xA = V_r^{\text{T}}(A) \ltimes x^{\text{T}}. \tag{11.3.1}$$

(2) 如果$Y \in M_{p\times n}$, 则

$$YA = (I_p \otimes V_r^{\text{T}}(A)) \ltimes V_r(Y). \tag{11.3.2}$$

证明 直接计算可以得到

$$\begin{aligned} xA &= V_r^{\text{T}}(A)x^{\text{T}} \\ &= \left(\sum_{i=1}^n a_{i1}x_i, \cdots, \sum_{i=1}^n a_{im}x_i\right), \end{aligned}$$

利用(11.3.1) 有

$$YA := \begin{bmatrix} Y^1 \\ \vdots \\ Y^p \end{bmatrix} A = \begin{bmatrix} V_r^{\mathrm{T}}(A)(Y^1)^{\mathrm{T}} \\ \vdots \\ V_r^{\mathrm{T}}(A)(Y^p)^{\mathrm{T}} \end{bmatrix} = (I_p \otimes V_r^{\mathrm{T}}(A))V_r(Y).$$

□

现在我们按如下方式展开$(A+BK)^t$:

$$(A+BK)^t = \sum_{i=0}^{2^t-1} P_i(A, BK),$$

其中P_i 是t 个元素的乘积, 且这些元素都是A 或BK. 它可以按如下方式构造: 将i 转换为长度为t 的二进制数, 并用“A” 替换“0”, “BK” 替换“1”, 再合并“K” 的指数相同的项, 这样我们就可以得到下面的表达式:

$$c_k(A+BK)^t B = \sum_{i=0}^{t}\sum_{j=1}^{T_i} S_0^{ij} K S_1^{ij} K \cdots S_{t-1}^{ij} K S_t^{ij}, \quad k = 1, \cdots, p. \tag{11.3.3}$$

其中$T_i = \begin{bmatrix} i \\ t \end{bmatrix}$. 利用引理11.3.1 和(11.3.1), (11.3.3) 可以表示成

$$\begin{aligned} & c_k(A+BK)^t B \\ &= \sum_{i=0}^{t}\sum_{j=1}^{T_i} S_0^{ij} \ltimes (I_m \otimes V_r^{\mathrm{T}}(S_1^{ij})) \ltimes Z \ltimes \cdots \ltimes (V_r^{\mathrm{T}}(I_m \otimes S_t^{ij})) \ltimes Z \\ &= \sum_{i=0}^{t}\sum_{j=1}^{T_i} S_0^{ij} \ltimes (I_m \otimes V_r^{\mathrm{T}}(S_1^{ij})) \ltimes (I_{m^2n} \otimes V_r^{\mathrm{T}}(S_2^{ij})) \\ &\qquad \ltimes \cdots \ltimes (I_{m^t n^{t-1}} \otimes V_r^{\mathrm{T}}(S_t^{ij})) \ltimes Z^t, \quad k = 1, \cdots, p. \end{aligned} \tag{11.3.4}$$

利用(11.3.4), 我们可以将$W(K)$ 和$T(K)$ 表示成标准的多项式形式如下:

$$\begin{aligned} & W(K) = W_0 + W_1 \ltimes Z + \cdots + W_{l-1} \ltimes Z^{l-1} \in M_{d\times m}, \quad d = \sum_{i=1}^{p} \rho_i - p. \\ & T(K) = T_0 + T_1 \ltimes Z + \cdots + T_l \ltimes Z^l \in M_{p\times m}. \end{aligned} \tag{11.3.5}$$

其中$l = \max\{\rho_i - 1 \mid i = 1, \cdots, p\}$.

设W^s 是由$W(K)$ 的s 行组成的集合, 则W^s 的大小是

$$|W^s| = \frac{d!}{s!(d-s)!}.$$

现在Morgan 问题可以表述成如下形式.

命题 11.3.2 Morgan 问题可解, 当且仅当存在$1 \le s \le m-p+1$ 使得

$$R(Z) := \sum_{L \in W^s} \det\left(L(Z)L^{\mathrm{T}}(Z)\right) = 0 \tag{11.3.6}$$

和

$$J(Z) := \sum_{L \in W^{s-1}} \det\left(\begin{bmatrix} T(Z) \\ L(Z) \end{bmatrix} \left(T^{\mathrm{T}}(Z)\ L^{\mathrm{T}}(Z)\right)\right) > 0 \tag{11.3.7}$$

有解Z.

证明 如果方程组(11.3.6)、(11.3.7) 有解$Z = V_r(K)$, 则由(11.3.6) 可得$\mathrm{rank}(W(K)) < s$, 且由(11.3.7) 可得

$$\mathrm{rank}\begin{pmatrix} T(K) \\ W(K) \end{pmatrix} = p+s-1.$$

于是根据定理11.2.4, 结论成立. □

现在Morgan 问题变为一个数值问题: 对于固定的$1 \le \rho_i \le n$, $i = 1, \cdots, p$ 和$1 \le s \le m-p+1$, 解方程组(11.3.6)、(11.3.7). 由于只有有限种情形, 因此只要这个问题在任何一种情形下可解, 则Morgan 问题可解. 有许多方法可以用来解这个数值问题.

例如, 我们可以将这个数值问题转化为"吴问题" [113]: 多项式$R(z) = 0$ 是否意味着$J(z) = 0$? 如果对所有情形都回答"是",则意味着Morgan 问题不可解, 否则只要对其中一种情形回答"否",则意味着Morgan 问题可解. 吴方法可以用来解决这个问题.

我们还可以将这个问题表述成一个优化问题:

$$\max_{R(z)=0} J(z).$$

如果最大值是零, 则Morgan 问题不可解, 否则可解.

注意到如果方阵$A(Z)$ 的每个元素都可以表示成形如$a_0 + a_1 \ltimes Z + \cdots + a_L \ltimes Z^L$ 的多项式，可以直接计算多项式形式的$\det(A(Z))$. 因此，为了得到(11.3.6)和(11.3.7) 我们只需要计算下面的乘积. 设

$$A = A_0 + A_1 \ltimes z + \cdots + A_s \ltimes z^s \in M_{m\times n},$$

$$B = B_0 + B_1 \ltimes z + \cdots + B_t \ltimes z^t \in M_{p\times n}.$$

于是

$$AB^{\mathrm{T}} = \sum_{i=0}^{s}\sum_{j=0}^{t} A_i \ltimes z^i \ltimes (z^{\mathrm{T}})^j \ltimes B_j^{\mathrm{T}}.$$

利用(11.3.1), 我们有

$$(z^{\mathrm{T}})^j = (z^{\mathrm{T}})^j I_{n^j} = V_r^{\mathrm{T}}(I_{n^j}) \ltimes z^j.$$

$$z^i \ltimes V_r^{\mathrm{T}}(I_{n^j}) = (I_{n^i} \ltimes V_r^{\mathrm{T}}(I_{n^j})) \ltimes z^i,$$

$$z^{i+j} \ltimes B_j^{\mathrm{T}} = (I_{n^{i+j}} \ltimes B_j^{\mathrm{T}}) \ltimes z^{i+j}.$$

利用它们, 我们有

$$AB^{\mathrm{T}} = \sum_{i=0}^{s}\sum_{j=0}^{t} A_i \ltimes \left(I_{n^i} \ltimes V_r^{\mathrm{T}}(I_{n^j})\right) \ltimes \left(I_{n^{i+j}} \ltimes B_j^{\mathrm{T}}\right) \ltimes Z^{i+j}. \tag{11.3.8}$$

显然当$1 \le \rho_i \le 2, i = 1, \cdots, p$ 时, 根据上面的推导, 我们有下面的命题

推论 11.3.3　对于$1 \le \rho_i \le 2, i = 1, \cdots, p$, 并设A1 成立, 则Morgan 问题可解, 当且仅当

$$J(Z) := \det\left(T(Z)T^{\mathrm{T}}(Z)\right) > 0 \tag{11.3.9}$$

有解Z.

此时Morgan 问题可以转化为无约束优化问题:

$$\max J(z).$$

如果最大值是零, 则Morgan 问题不可解, 否则可解.

综上所述, 我们给出解Morgan 问题的如下步骤:

(1) 对于$\rho_1, \cdots, \rho_p = 1, \cdots, n$, 利用(11.3.3)~(11.3.4) 将$T(K)$ 和$W(K)$ 表示成标准的多项式形式(11.3.5).

(2) 对于$s=1,\cdots,m-p+1$ 和每个$L\in W^s$, 利用(11.3.8) 计算出

$$L(z)L^{\mathrm{T}}(z),\quad \begin{bmatrix} T(z)\\ L(z)\end{bmatrix}\begin{bmatrix} T^{\mathrm{T}}(z) & L^{\mathrm{T}}(z)\end{bmatrix}.$$

(3) 分别用(11.3.6) 和(11.3.7) 计算出$R(z)$ 和$J(z)$.

(4) 解数值问题

$$\begin{cases} R(z)=0,\\ J(z)>0.\end{cases} \tag{11.3.10}$$

其中$R(z)$ 和$J(z)$ 都是多项式.

例 11.3.4 考虑线性系统

$$\begin{cases} \dot{x}=\begin{bmatrix} 0 & 6 & -2 & -4 & 0\\ -1 & 0 & 0 & 0 & 1\\ 0 & -2 & 1 & 1 & -1\\ -1 & 1 & -1 & -1 & 1\\ 0 & 2 & 0 & 0 & 1\end{bmatrix}x+\begin{bmatrix} -2 & 1 & 0\\ 0 & 0 & 0\\ 2 & 0 & -2\\ -1 & 0 & 1\\ -1 & 1 & 1\end{bmatrix}u,\\ y=\begin{bmatrix} 0 & 1 & -1 & -1 & 0\\ 0 & 1 & 0 & -1 & 0\end{bmatrix}x.\end{cases} \tag{11.3.11}$$

注意到$\rho_1+\rho_2$ 不能大于状态的维数5 (请读者自行证明这一点), 我们要检验的可能的情形是$\rho_1=1,\rho_2=1,2,3,4;\rho_1=2,\rho_2=1,2,3;\rho_1=3,\rho_2=1,2;\rho_1=4,\rho_2=1$. 作为例子我们只检验$\rho_1=3,\rho_2=2$ 的情形, 此时

$$W(K)=\begin{bmatrix} c_1B\\ c_1(A+BK)B\\ c_2B\end{bmatrix}=\begin{bmatrix} -1 & 0 & 1\\ p_1(Z) & p_2(Z) & p_3(Z)\\ 1 & 0 & -1\end{bmatrix},$$

这里

$$Z=V_r(K)=(k_{11},\ldots,k_{15},\ \ldots\ ,k_{31},\ldots,k_{35})^{\mathrm{T}}.$$

并且利用(4.1.32) 有

$$\begin{bmatrix} p_1(Z)\\ p_2(Z)\\ p_3(Z)\end{bmatrix}=V_r(c_1AB+c_1BKB)=\begin{bmatrix} -1\\ 0\\ 2\end{bmatrix}+((c_1B)\otimes B^{\mathrm{T}})Z.$$

可以算出

$$
\begin{aligned}
p_1(Z) &= -1 + 2k_{11} - 2k_{13} + k_{14} + k_{15} - 2k_{31} + 2k_{33} - k_{34} - k_{35}, \\
p_2(Z) &= 1 - k_{11} - k_{15} + k_{31} + k_{35}, \\
p_3(Z) &= 1 - k_{11} + 2k_{13} - k_{14} - 2k_{15} + k_{31} - 2k_{33} + k_{34} + 2k_{35}.
\end{aligned}
$$

另外

$$
\begin{aligned}
T(K) &= \begin{bmatrix} c_1(A+BK)^2B \\ c_2(A+BK)B \end{bmatrix} \\
&= \begin{bmatrix} c_1A^2B \\ c_2AB \end{bmatrix} + \begin{bmatrix} c_1ABKB + c_1BKAB \\ c_2BKB \end{bmatrix} + \begin{bmatrix} c_1BKBKB \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -1 \end{bmatrix} + T_1Z + T_2Z^2,
\end{aligned}
$$

这里

$$
T_1 = \left[\begin{matrix} 2 & -1 & 0 & -1 & 0 & -1 & -4 & 1 & 4 & 1 & 0 & -3 & 2 & -2 & -2 \\ -2 & 1 & 0 & 0 & 0 & 0 & 2 & 0 & -2 & -1 & 0 & 1 & -1 & 1 & 1 \end{matrix}\right.
$$

$$
\begin{matrix} -2 & 1 & 0 & 0 & 0 & 0 & 2 & 0 & -2 & -1 & 0 & 1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{matrix}
$$

$$
\left.\begin{matrix} -2 & 1 & 0 & 1 & 0 & 1 & 4 & -1 & -4 & -1 & 0 & 3 & -2 & 2 & 2 \\ 2 & -1 & 0 & 0 & 0 & 0 & -2 & 0 & 2 & 1 & 0 & -1 & 1 & -1 & -1 \end{matrix}\right].
$$

T_2 是一个2×675 的矩阵, 由于篇幅关系不在这里写出, 读者可自行计算.

根据命题11.3.2, 我们需要对$s = 1$ 和$s = 2$ 两种情况进行检验, 显然$s = 1$ 时$R(Z) > 0$, 对于$s = 2$, 由(11.2.5) 可知, 欲使D 非奇异, $\text{rank}(H) \geq 2$, 于是由(11.2.4), $\text{rank}\, W(K) \leq 1$, 因此不妨设

$$
p_1(Z) = 0, \quad p_2(Z) = 0, \quad p_3(Z) = 0. \tag{11.3.12}
$$

容易看出其一组解为

$$
K = V_r^{-1}(Z) = \begin{bmatrix} 0 & 2 & -1 & -2 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 2 & -1 & -2 & -1 \end{bmatrix},
$$

此时

$$W(K)=\begin{bmatrix}-1&0&1\\0&0&0\\1&0&-1\end{bmatrix},\qquad T(K)=\begin{bmatrix}1&0&1\\0&1&0\end{bmatrix}.$$

显然$R(Z)=0$, 并且

$$\begin{aligned}J(Z)\ &=\ det\left(\begin{bmatrix}1&0&-1\\0&1&0\\-1&0&1\end{bmatrix}\begin{bmatrix}1&0&-1\\0&1&0\\-1&0&1\end{bmatrix}^{\mathrm{T}}\right)\\&\quad+det\left(\begin{bmatrix}1&0&-1\\0&1&0\\0&0&0\end{bmatrix}\begin{bmatrix}1&0&-1\\0&1&0\\0&0&0\end{bmatrix}^{\mathrm{T}}\right)\\&\quad+det\left(\begin{bmatrix}1&0&-1\\0&1&0\\1&0&-1\end{bmatrix}\begin{bmatrix}1&0&-1\\0&1&0\\1&0&-1\end{bmatrix}^{\mathrm{T}}\right)\\&=\ 4>0.\end{aligned}$$

于是根据命题11.3.2 可知系统(11.3.11) 的Morgan 问题可解. 下面我们寻找反馈矩阵H, 因为$\mathrm{Span}\{\mathrm{col}(H)\}\subset\ker(W(K))$, 不妨选

$$H=\begin{bmatrix}0&1\\1&0\\0&1\end{bmatrix},$$

于是解耦控制为$u=Kx+Hv$, 反馈系统为

$$\begin{cases}\dot{x}=\begin{bmatrix}0&1&0&0&0\\-1&0&0&0&1\\0&-2&1&1&1\\-1&1&-1&-1&0\\0&1&0&0&0\end{bmatrix}x+\begin{bmatrix}1&1\\0&0\\0&-2\\0&1\\1&1\end{bmatrix}v,\\ y=\begin{bmatrix}0&1&-1&-1&0\\0&1&0&-1&0\end{bmatrix}x.\end{cases}\tag{11.3.13}$$

我们可以检验

$$WH=\begin{bmatrix}0&0\\0&0\\0&0\end{bmatrix},\quad D=TH=\begin{bmatrix}0&2\\1&0\end{bmatrix}.$$

于是我们有$\rho_1=3$, $\rho_2=2$, 并且解耦阵D 非奇异. □

注 (1) 从上例讨论可知满足(11.3.12) 的所有K 均为Morgan 问题的可行解.

(2) 利用上例中的K, H 可以检验引理11.2.1 中的(11.2.1) 成立, (11.2.2) 非奇异, 引理11.2.3 中的(11.2.6) 成立, $T(K)$ 行满秩.

(3) 对于定理11.2.4, 易证

$$\operatorname{rank}\left(\begin{bmatrix}T(K_0)\\W(K_0)\end{bmatrix}\right)=3,\quad \operatorname{rank}(W(K_0))=1,\quad p=2.$$

故(11.2.7) 成立.

11.4 注释与参考

输入输出解耦问题在控制论中的重要性是显见的, 举机器人为例, 如果不能实现解耦, 那么某个控制, 在它让机器人举手时, 同时它也让机器人抬腿, 这就成了笑话了. 这个问题首先由Morgan 在1964 年提出, 因此,后来就把它称为“Morgan 问题”. 近半个世记以来有大量的文章研究这个问题. 第一个重要的进展是$m=p$ 的情形, 这种情形由Falb 和Wolovich 完全解决[47, 110], 他们的结果也是我们的方法的出发点.

此后对Morgan 问题的努力虽有许多进展, 但都无法突破$m>p$ 的情形. 曾经有几次宣布解决了这个问题, 但是随后不久发现的反例都推翻了它们, 这些可参见文献[112], [94], [42], [60], 其中文献[111] 对Morgan 问题给出了一个系统的描述. 不幸地是, 经过40余年Morgan 问题仍然是一个公开的难题.

非线性系统的Morgan 问题也同样是极具挑战性的问题. 当$m>p$ 时, 在文献[41] 中讨论了利用动态补偿器来解决这个问题, 并且得到了一个充分和必要的条件, 同时它也是利用静态状态反馈可解的必要条件. 一些后续进展可在文献[53], [44] 中找到.

第 12 章 非线性系统的线性化

由于线性系统相对简单, 并且已有许多成熟的分析和综合工具, 于是线性化, 即将一个非线性系统精确或近似地转化成线性系统, 就成了数学、物理和控制中的一个重要课题. 通过合适的坐标变换(对于控制系统可能包括合适的控制) 将一个非线性动态系统表示成线性形式称为线性化. 本章主要考虑一些特殊的线性化问题, 半张量积在其中起着关键作用.

12.1 Carleman 线性化

考虑动态系统

$$\dot{x} = f(x), \quad x \in \mathbb{R}^n, \tag{12.1.1}$$

其中$f(x)$ 是解析向量场, 且$f(0) = 0$.

J. Carleman 提出了一个方法将它嵌入到一个无限维线性系统中. 本节将介绍其基本形式及实现. 它的原始形式比较复杂, 半张量积方法使它得到简化.

选择x, x^2, $\cdots$ 作为一组基底, 则系统(12.1.1) 可以表示成

$$\dot{x} = F_1 x + F_2 x^2 + F_3 x^3 + \cdots, \tag{12.1.2}$$

其中F_1 是一个$n \times n$ 矩阵, F_2 是一个$n \times n^2$ 矩阵, 等等.

我们可以考虑x^2, x^3 等作为一组独立的变量, 然后计算它们的导数来得到Carleman 线性化形式如下:

$$\begin{bmatrix} \dot{x} \\ \dot{x^2} \\ \dot{x^3} \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} & A_{14} & \cdots \\ 0 & A_{22} & A_{23} & A_{24} & \cdots \\ 0 & 0 & A_{33} & A_{34} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} x \\ x^2 \\ x^3 \\ \vdots \end{bmatrix}. \tag{12.1.3}$$

定理 12.1.1 在Carleman 线性化形式(12.1.3) 中, 系数A_{ij} 由下面的方程决定.

$$\begin{cases} A_{1i} = F_i, \quad i \geq 1, \\ A_{k,k+s} = \sum_{i=0}^{k-1} I_{n^i} \otimes F_{s+1} \otimes I_{n^{k-1-i}}. \end{cases} \tag{12.1.4}$$

证明　根据链式法则, 我们有

$$\frac{d}{dt}(x^k)=\sum_{i=0}^{k-1}x^i\dot{x}x^{k-i-1}=\sum_{s=0}^{\infty}\sum_{i=0}^{k-1}x^iF_{s+1}x^{k-i+s}.$$

利用(3.3.21), 有

$$x^iF_{s+1}x^{k-i+s}=(I_{n^i}\otimes F_{s+1})\ltimes x^{k+s}=(I_{n^i}\otimes F_{s+1}\otimes I_{n^{k-i-1}})x^{k+s}.$$

(12.1.4) 成立. □

我们可以将(12.1.3) 表示成如下线性形式

$$\dot{X}=AX, \tag{12.1.5}$$

其中A 是一个无限维块上三角矩阵.

无限维块上三角矩阵具有一些特殊的性质, 它使表达式(12.1.5) 有意义. 这里我们给出一个简要的讨论:

记无限维块上三角矩阵A 的左上k 个块子式为A_k, 即

$$A_k=\begin{bmatrix}A_{11} & A_{12} & \cdots & A_{1k}\\ 0 & A_{22} & \cdots & A_{2k}\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & A_{kk}\end{bmatrix}.$$

我们说具有方对角块的块上三角矩阵具有一组结构常数$(k_1,k_2,\cdots)$, 如果对角块的维数为$\dim(A_{ii})=k_i\times k_i$. 例如, 对于式(12.1.3), 它的结构常数为$(n,n^2,n^3,\cdots)$. 为了叙述简便, 我们将A_k 等同于它的推广: 无限维块上三角矩阵A_k^e, 且以A_k 作为它的左上子式, 其余元素为0. 利用这个约定, 系数矩阵A 可以看成

$$A=\lim_{k\to\infty}A_k.$$

这个极限是有定义的, 因为如果我们记A_k 的第(ij) 个元素为a_{ij}^k, 那么序列$\{a_{ij}^k,\ k=1,2,\cdots\}$ 具有如下形式

$$(a_{ij}^1,a_{ij}^2,\cdots,a_{ij}^k,\cdots)=(0,\cdots,0,c_{ij},c_{ij},c_{ij},\cdots).$$

即除了开始的有限项以外它是一个常数序列. 基于相同的道理, 下面的记号也有定义.

定义 12.1.2 (1) 设A 和B 是两个具有相同结构常数的无限维块上三角矩阵. 我们可以定义A 和B 的乘积为

$$AB := \lim_{k\to\infty} A_k B_k.$$

(2) 假设A_{ii}, $i=1,2,\cdots$ 可逆, 则我们可定义

$$A^{-1} := \lim_{k\to\infty} A_k^{-1}.$$

(3)

$$e^A := \lim_{k\to\infty} e^{A_k}.$$

现在很自然地可以将线性解

$$X = e^{At}X_0$$

作为非线性系统(12.1.3) 的解. 实际上, 我们可以仅用有限项来近似这个解.

记$e^{A_k t}$ 的第(i,j) 个块为$E_{ij}^k(t)$, 则容易看出

$$E_{ij}^s(t) = E_{ij}^k(t), \quad s>k, \quad i,j\le k.$$

因此我们可以定义

$$X^n(t) = \sum_{k=1}^n E_{1k}^k(t)X_0^k.$$

由(12.1.3), 显然如果

$$X(t) = \lim_{n\to\infty} X^n(t)$$

存在, 那么它是以$X(0)=X_0$ 为初值的(12.1.1) 的解.

我们特别对由Carleman 线性化得到的上三角矩阵感兴趣. 在Carleman 线性化形式(12.1.3) 中假设$F_1 = A_{11}$ 是稳定的(反稳定的) , 即$\mathrm{Re}\,\sigma(A_{11})<0$ ($\mathrm{Re}\,\sigma(A_{11})>0$), 那么$A$ 是可逆的. 实际上, 我们有

定理 12.1.3 假设$F_1 = A_{11}$ 具有特征值$\sigma(A_{11}) = \{\lambda_1,\cdots,\lambda_n\}$, 那么$A_{ii}$, $i\ge 2$ 有特征值

$$\sigma(A_{ii}) = \{\lambda_{k_1}+\cdots+\lambda_{k_i} \mid k_1,\cdots,k_i = 1,\cdots,n\}.$$

证明 首先我们假设A_{11} 的特征值互不相同, 并且相应的特征向量为

$$\{\xi_1,\cdots,\xi_n\}.$$

那么直接计算有

$$A_{ii}(\xi_{k_1} \ltimes \cdots \ltimes \xi_{k_i}) = (\lambda_{k_1} + \cdots + \lambda_{k_i})(\xi_{k_1} \ltimes \cdots \ltimes \xi_{k_i}).$$

为了避免记号混乱，我们仅说明$i = 2$ 时的情形. 通过(12.1.4) 我们有

$$A_{22} = I_n \otimes A_{11} + A_{11} \otimes I_n.$$

于是

$$\begin{aligned} A_{22}(\xi_i \ltimes \xi_j) &= (I_n \otimes A_{11} + A_{11} \otimes I_n)(\xi_i \ltimes \xi_j) \\ &= (I_n \otimes A_{11})(\xi_i \ltimes \xi_j) + (A_{11} \otimes I_n)(\xi_i \ltimes \xi_j) \\ &= \lambda_j \xi_i \ltimes \xi_j + \lambda_i \xi_i \ltimes \xi_j. \end{aligned}$$

既然所有的λ_i 互不相同，我们可以断言所有的

$$\{\xi_i \ltimes \xi_j \mid i, j = 1, \cdots, n\}$$

线性无关. 实际上，由于λ_i 互不相同，所以所有的ξ_i 线性无关. 直接计算有

$$(\xi_1, \cdots, \xi_n) \ltimes (\xi_1, \cdots, \xi_n) = (\xi_1\xi_1, \cdots, \xi_1\xi_n, \cdots, \xi_n\xi_1, \cdots, \xi_n\xi_n)$$

断言成立.

因此A_{22} 的特征值为

$$\sigma(A_{22}) = \{\lambda_i + \lambda_j \mid i, j = 1, \cdots, n\}.$$

由连续性，只要多重特征值存在，特征值结构也是对的. □

下面，我们考虑多项式系统

$$\dot{x} = F_1 x + F_2 x^2 + \cdots + F_k x^k. \tag{12.1.6}$$

Carleman 线性化技巧可以用来研究形式

$$H(t, x) = e^{-\xi t} P(x)$$

的一次积分.

向量场$f(x)$ 的一次积分指一(时变) 函数$\phi(t, x)$ 满足

$$\frac{d}{dt}\phi(t, x) = \frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial x} f(x) = 0.$$

假设$P(x) = P_0 + P_1 x + \cdots + P_s x^s$, 它的对称系数为$P_1, \cdots, P_s$. (这里的"对称"意味着具有不同因子次序的相同单项式的系数是一样的. 例如, $x_1^2 x_2$, $x_1 x_2 x_1$ 和$x_2 x_1^2$ 的系数是一样的.) 容易看出, 如果$\xi \neq 0$, 则$P_0 = 0$. 因此, 我们可以简单地假设$P_0 = 0$.

令$dH(t,x)/dt = 0$, 我们有

$$\begin{aligned}&\begin{bmatrix} P_1 & \cdots & P_s \end{bmatrix} \begin{bmatrix} A_{11} & \cdots & \cdots & A_{1k} & & & \\ & \ddots & & & \ddots & & \\ & & A_{ss} & \cdots & \cdots & A_{s,s+k-1} \end{bmatrix} \begin{bmatrix} x \\ x^2 \\ \vdots \\ x^{s+k-1} \end{bmatrix} \\ &= \xi \begin{bmatrix} P_1 & \cdots & P_s \end{bmatrix} \begin{bmatrix} x \\ x^2 \\ \vdots \\ x^s \end{bmatrix}. \end{aligned} \tag{12.1.7}$$

由于基底x^k 是冗余的, 因此系数不唯一. 利用这种形式查找一次积分太保守. 另外, 得到的条件一般也不是必要的, 因为在其他等价系数下也可以得到其他类型的一次积分.

为了得到充要条件, 我们要将系统转换到自然基底下. 令

$$P_i = \tilde{P}_i \mathrm{T}_B(n,i), \quad \tilde{A}_{ij} = \mathrm{T}_B(n,i) A_{ij} \mathrm{T}_N(n,j).$$

将它们代入到(12.1.7) 得到

$$\begin{aligned}&\begin{bmatrix} \tilde{P}_1 & \cdots & \tilde{P}_s \end{bmatrix} \begin{bmatrix} \tilde{A}_{11} & \cdots & \cdots & \tilde{A}_{1k} & & & \\ & \ddots & & & \ddots & & \\ & & \tilde{A}_{ss} & \cdots & \cdots & \tilde{A}_{s,s+k-1} \end{bmatrix} \begin{bmatrix} x \\ x_{(2)} \\ \vdots \\ x_{(s+r-1)} \end{bmatrix} \\ &= \xi \begin{bmatrix} \tilde{P}_1 & \cdots & \tilde{P}_s \end{bmatrix} \begin{bmatrix} x \\ x_{(2)} \\ \vdots \\ x_{(s)} \end{bmatrix}. \end{aligned}$$

定理 12.1.4 记$h_i = \tilde{P}_i^{\mathrm{T}}$, $B_{ij} = \tilde{A}_{ji}^{\mathrm{T}}$, 则系统(12.1.6) 具有一次积分$H(t,x)$,

当且仅当, 存在ξ 使得下面的方程组具有非零解$(h_1,\cdots,h_s)$.

$$\begin{cases}\begin{bmatrix}B_{11} & 0 & 0 & \cdots & 0\\ B_{21} & B_{22} & 0 & \cdots & 0\\ \vdots & & & & \\ B_{s1} & B_{s2} & & \cdots & B_{s,s}\end{bmatrix}\begin{bmatrix}h_1\\ h_2\\ \vdots\\ h_s\end{bmatrix}=\xi\begin{bmatrix}h_1\\ h_2\\ \vdots\\ h_s\end{bmatrix},\\ \begin{bmatrix}B_{s+1,1} & B_{s+1,2} & \cdots & B_{s+1,s}\\ \vdots & & & \\ B_{k,1} & B_{k,2} & \cdots & B_{k,s}\\ 0 & B_{k+1,2} & \cdots & B_{k+1,s}\\ \vdots & & & \\ 0 & \cdots & 0 & B_{s+k-1,s}\end{bmatrix}\begin{bmatrix}h_1\\ h_2\\ \vdots\\ h_s\end{bmatrix}=0.\end{cases} \tag{12.1.8}$$

接着, 我们考虑方程组(12.1.8) 的解的形式.

下面的引理本身也很有趣. 设$I=(i_1,\cdots,i_k)$, 我们用$Id(I;n^k)$ 代替$Id(i_1,\cdots,i_k;n,\cdots,n)$. (关于索引可参见第1 章.)

引理 12.1.5 假设行向量$h\in\mathbb{R}^{n^k}$ 按照索引$Id(I;n^k)$ 排列, 并且它关于$Id(I;n^k)$ 对称, $F\in M_{n\times n}$. 令

$$A=F\otimes I_{n^{k-1}}+I_n\otimes F\otimes I_{n^{k-2}}+\cdots+I_{n^{k-1}}\otimes F.$$

则hA 也关于$Id(I;n^k)$ 对称.

证明 由于交换任意两个指标都可以通过交换相邻的指标来实现, 因此我们只需说明在任意两个相邻指标的交换下hA 是不变的. 令

$$\Phi=I_{n^{j-1}}\otimes W_{[n]}\otimes I_{n^{k-j-1}}.$$

显然交换h 的第j 个和第$(j+1)$ 个指标得到新向量$h\Phi$. 由于h 关于$Id(I;n^k)$ 对称, 那么$h\Phi=h$, 并且$h\Phi A=hA$. 为了看到hA 对称, 即$hA\Phi=hA$, 只要证明

$$A\Phi=\Phi A$$

就行了. 注意到A 中的项具有下面的形式

$$I\otimes\cdots\otimes I\otimes F\otimes I\otimes\cdots\otimes I. \tag{12.1.9}$$

如果F 不位于第j 或$j+1$ 位, 那么显然Φ 同(12.1.9) 可交换. 因此我们只需考虑A 中有关的两项. 这就说明

$$W_{[n]}(F\otimes I+I\otimes F)=(F\otimes I+I\otimes F)W_{[n]}.$$

注意到$W_{[n]}^{-1}=W_{[n]}$ (参见第1 章(1.5.2)), 意味着上面的等式成立. □

命题 12.1.6 设(12.1.6) 中F_1 的特征值为$\sigma=\{\lambda_1,\cdots,\lambda_n\}$, 则(12.1.8) 中$B_{kk}$ 的特征值为

$$\sigma(B_{kk})=\{\lambda_{i_1}+\cdots+\lambda_{i_k}\,|\,i_1,\cdots,i_k=1,\cdots,n\}.$$

证明 由于$B_{kk}=\tilde{A}_{kk}^{\mathrm{T}}$, 且$A_{kk}$ 的特征值为σ_k, 这就足够说明$\tilde{A}_{kk}$ 和A_{kk} 具有相同的特征值. 设μ 是$\tilde{A}_{kk}$ 的一个特征值, 那么存在$\tilde{P}\neq 0$ 使得

$$\tilde{P}\tilde{A}_{kk}=\mu\tilde{P}.$$

由定义有$\tilde{A}_{kk}=\mathrm{T}_B(n,k)A_{kk}\mathrm{T}_N(n,k)$. 注意到$\mathrm{T}_B(n,k)\mathrm{T}_N(n,k)=I$, 于是

$$\tilde{P}\mathrm{T}_B(n,k)A_{kk}\mathrm{T}_N(n,k)\begin{bmatrix}x_{(1)}\\x_{(2)}\\\vdots\\x_{(k)}\end{bmatrix}=\mu\tilde{P}\mathrm{T}_B(n,k)\mathrm{T}_N(n,k)\begin{bmatrix}x_{(1)}\\x_{(2)}\\\vdots\\x_{(k)}\end{bmatrix}. \tag{12.1.10}$$

令$P=\tilde{P}\mathrm{T}_B(n,k)$, 则$P\neq 0$ 是一个对称集合. 对于P, (12.1.10) 变成

$$PA_{kk}\begin{bmatrix}x^1\\x^2\\\vdots\\x^k\end{bmatrix}=\mu P\begin{bmatrix}x^1\\x^2\\\vdots\\x^k\end{bmatrix}. \tag{12.1.11}$$

根据引理12.1.5, PA_{kk} 仍是一个对称集合. 由对称系数的唯一性, 我们有

$$PA_{kk}=\mu P.$$

因此, μ 也是A_{kk} 的一个特征值.

反之, 假设μ 是A_{kk} 的一个特征值, 则$\mu=\lambda_{i_1}+\cdots+\lambda_{i_k}$. 记$Y_j$ 为F_1 关于λ_{i_j} 的特征向量, 于是我们构造

$$Y=\sum_{\sigma\in\mathbf{S}_k}Y_{\sigma(1)}\otimes\cdots\otimes Y_{\sigma(k)},$$

其中$\mathbf{S}_k$ 是k 阶置换群. 显然有

$$YA_{kk} = \mu Y.$$

由于Y 对称, 于是存在$\tilde{Y} \neq 0$ 使得$Y = \tilde{Y}\mathrm{T}_B(n,k)$, 这样得到

$$\tilde{Y}\mathrm{T}_B(n,k)A_{kk} = \mu\tilde{Y}\mathrm{T}_B(n,k).$$

上式两边同时右乘$\mathrm{T}_N(n,k)$ 得到

$$\tilde{Y}\tilde{A}_{kk} = \mu\tilde{Y}.$$

即μ 也是$\tilde{A}_{kk}$ 的特征值. □

命题 12.1.7　(1) 如果(12.1.8) 有解$h \neq 0$, 那么

$$\xi = c_1\lambda_{i_1} + \cdots + c_s\lambda_{i_s},$$

其中$\lambda_{i_1}, \cdots, \lambda_{i_s} \in \sigma(F_1)$; $c_1, \cdots, c_s$ 取1 或0.

(2) 如果h 有一个分量$h_j \neq 0$, 那么$\xi \in \sigma^j$. 如果h 有t 个非零分量, 那么σ^s 至少有一个t 重元素, 这里$\sigma^t = \{c_1\lambda_{i_1} + \cdots + c_t\lambda_{i_t} \,|\, c_1, \cdots, c_t \in \{0,1\}\}$.

(3) 如果(12.1.6) 有一个线性一次积分$H(t,x) = e^{-\xi t}h^{\mathrm{T}}x$, 那么对于任意整数$j > 0$, $H_j(t,x) = e^{-j\xi t}(h^{\mathrm{T}})^j x^j$ 是(12.1.6) 的一个一次积分.

证明　(1) 和(2) 是命题12.1.6 的直接推论. 我们证明(3). 如果(12.1.6) 有一个线性一次积分$H(t,x) = e^{-\xi t}h^{\mathrm{T}}X$, 那么

$$\begin{cases} F_1h = \xi h, \\ F_ih = 0, \quad i = 2, \cdots, k. \end{cases}$$

设$p = (0_n, 0_{n^2}, \cdots, 0_{n^{j-1}}, h^j)$, 其中$0_k$ 是$\mathbb{R}^k$ 中的零向量. 由于

$$A_{j,j+s-1} = I_{n^{j-1}} \otimes F_s + I_{n^{j-2}} \otimes F_s \otimes I + \cdots + F_s \otimes I_{n^{j-1}},$$

于是

$$\begin{cases} A_{jj}h^j = j\xi h^j, \\ A_{jt}h^j = 0, \quad t = j+1, \cdots, j+k-1. \end{cases}$$

这就意味着p 满足由$j\xi$ 取代ξ 的(12.1.7). □

命题12.1.7 提供了一个方便的工具来查找一次积分. 实际上, 固定ξ 后, 问题变为求解线性代数系统的问题. 对于Lorenz 系统, 命题12.1.7 中的(1) 和(2) 已在文献[12] 中提及. 因此在这里叙述的结果是对他们工作的推广.

例 12.1.8 Lotka-Volterra 方程建立了生物种群和化学反应相互作用的模型. Lotka-Volterra 方程可以写为

$$\dot{x}_i = x_i\left(a_i + \sum_{j=1}^{n} b_{ij}x_j\right), \quad i = 1, \cdots, n. \tag{12.1.12}$$

设$n = 2$. 令

$$A_{11} = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} b_{11} & b_{12} & 0 & 0 \\ 0 & 0 & b_{21} & b_{22} \end{bmatrix}.$$

于是

$$\begin{aligned} A_{22} &= A_{11} \otimes I_2 + I_2 \otimes A_{11} = \begin{bmatrix} 2a_1 & 0 & 0 & 0 \\ 0 & a_1 + a_2 & 0 & 0 \\ 0 & 0 & a_1 + a_2 & 0 \\ 0 & 0 & 0 & 2a_2 \end{bmatrix}, \\ A_{23} &= A_{12} \otimes I_2 + I_2 \otimes A_{12} \\ &= \begin{bmatrix} 2b_{11} & b_{12} & b_{12} & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{11} & b_{21} & b_{12} + b_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{11} + b_{21} & b_{12} & b_{22} & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{21} & b_{21} & 2b_{22} \end{bmatrix}. \end{aligned}$$

Carleman 线性化形式变成

$$\begin{bmatrix} \dot{x} \\ \dot{x}^2 \\ \dot{x}^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & 0 & 0 & \cdots \\ 0 & A_{22} & A_{23} & 0 & \cdots \\ 0 & 0 & A_{33} & A_{34} & \cdots \\ \vdots & & & & \end{bmatrix} \begin{bmatrix} x \\ x^2 \\ x^3 \\ \vdots \end{bmatrix}.$$

例如, 我们查找形式$H(t, x) = e^{-\xi t}(P_1 x + P_2 x^2)$ 的一次积分. 则

$$\mathrm{T}_B(2,2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathrm{T}_N(2,2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathrm{T}_N(2,3)=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&1&0&0\\0&0&1&0\\0&1&0&0\\0&0&1&0\\0&0&1&0\\0&0&0&1\end{bmatrix},$$

$$B_{11}=A_{11}^{\mathrm{T}},\quad B_{21}=\tilde{A}_{12}^{\mathrm{T}}=\mathrm{T}_B(2,1)A_{12}\mathrm{T}_N(2,2)=\begin{bmatrix}b_{11}&b_{12}&0\\0&b_{21}&b_{22}\end{bmatrix},$$

$$B_{22}^{\mathrm{T}}=\mathrm{T}_B(2,2)A_{22}\mathrm{T}_N(2,2)=\begin{bmatrix}2a_1&0&0\\0&a_1+a_2&0\\0&0&2a_2\end{bmatrix},$$

$$B_{32}^{\mathrm{T}}=\mathrm{T}_B(2,2)A_{23}\mathrm{T}_N(2,3)=\begin{bmatrix}2b_1&2b_{12}&0&0\\0&b_{11}+b_{21}&b_{12}+b_{22}&0\\0&0&2b_{21}&2b_{22}\end{bmatrix}.$$

我们可以得出结论, 二阶的一次积分存在, 当且仅当下面的方程具有非零解.

$$\begin{cases}\begin{bmatrix}B_{11}&0\\B_{21}&B_{22}\end{bmatrix}\begin{bmatrix}h_1\\h_2\end{bmatrix}=\xi\begin{bmatrix}h_1\\h_2\end{bmatrix},\\ B_{32}h_2=0.\end{cases}$$

其中只可能$\xi\in\{a_1,a_2,2a_1,2a_2,a_1+a_2\}$. □

12.2 平面多项式系统的不变量

本节我们考虑平面多项式系统的非多项式不变量. 考虑多项式系统

$$\dot{x}=F_0+F_1x+\cdots+F_kx^k,\quad x\in\mathbb{R}^2. \tag{12.2.1}$$

考虑的不变量具有如下形式

$$H(x,t)=e^{\xi t}x_1^{\alpha}x_2^{\beta}(P_0+P_1x+\cdots+P_lx^l). \tag{12.2.2}$$

这种形式的不变量在文献[12] 中已有研究. 近来, Darboux 方法被用来查找这样的不变量[13].

我们的目的是将这个问题转化为一组代数方程. 利用(5.5.4), $H(x,t)$ 关于时间的导数是

$$\begin{aligned}\tfrac{dH}{dt} &= \tfrac{\partial H}{\partial t} + DH\cdot\dot{x}\\ &= \xi e^{\xi t}x_1^{\alpha}x_2^{\beta}(P_0+P_1x+\cdots+P_lx^l)\\ &\quad+e^{\xi t}(\alpha x_1^{\alpha-1}x_2^{\beta},\ \beta x_1^{\alpha}x_2^{\beta-1})(P_0+P_1x+\cdots+P_lx^l)(F_0+F_1x+\cdots+F_kx^k)\\ &\quad+e^{\xi t}x_1^{\alpha}x_2^{\beta}(P_1+P_2\Phi_1x+\cdots+P_l\Phi_{l-1}x^{l-1})(F_0+F_1x+\cdots+F_kx^k).\end{aligned}$$

令

$$\frac{dH}{dt}=0,$$

并注意到

$$x_1x_2=(0\ 1\ 0\ 0)x^2,\quad (\alpha x_2,\ \beta x_1)=(0\ \beta\ \alpha\ 0)x.$$

我们有

$$\begin{aligned}&\xi(0\ 1\ 0\ 0)x^2(P_0+P_1x+\cdots+P_lx^l)\\ &\quad+(0\ \beta\ \alpha\ 0)x(P_0+P_1x+\cdots+P_lx^l)(F_0+F_1x+\cdots+F_kx^k)\\ &\quad+(0\ 1\ 0\ 0)x^2(P_1+P_2\Phi_1x+\cdots+P_l\Phi_{l-1}x^{l-1})(F_0+F_1x+\cdots+F_kx^k)\\ &=0.\end{aligned}$$

利用(3.3.21), 它可以表示成

$$\begin{aligned}&\xi(0\ 1\ 0\ 0)[(I_4\otimes P_0)x^2+(I_4\otimes P_1)x^3+\cdots+(I_4\otimes P_l)x^{l+2}]\\ &\quad+(0\ \beta\ \alpha\ 0)[(I_2\otimes P_0)x+(I_2\otimes P_1)x^2\\ &\quad+\cdots+(I_2\otimes P_l)x^{l+1}](F_0+F_1x+\cdots+F_kx^k)\\ &\quad+(0\ 1\ 0\ 0)[(I_4\otimes P_1)x^2+(I_4\otimes P_2\Phi_1)x^3\\ &\quad+\cdots+(I_4\otimes P_l\Phi_{l-1})x^{l+1})(F_0+F_1x+\cdots+F_kx^k)\\ &=0.\end{aligned}$$

再次利用(3.3.21), 将乘积乘出来得到

$$
\begin{aligned}
&\sum_{s=1}^{l+1}\xi(0\ 1\ 0\ 0)[I_4\otimes P_{s-1}]x^{s+1}\\
&\quad+\sum_{s=0}^{k+l}(0\ \beta\ \alpha\ 0)\sum_{i=0,j=0}^{i+j=s}[(I_2\otimes P_i)(I_{2^{i+1}}\otimes F_{s-i})]x^{s+1}\\
&\quad+\sum_{s=1}^{k+l}(0\ 1\ 0\ 0)\sum_{i=1,j=0}^{i+j=s}[(I_4\otimes P_i\Phi_{i-1})(I_{2^{i+1}}\otimes F_{s-i})]x^{s+1}\\
&=0.
\end{aligned}
$$

逐项转化为自然基底, 我们有下面的结果.

定理 12.2.1 系统(12.2.1) 具有(12.2.2) 形式的不变量, 当且仅当下面的代数方程组有解$(\xi,\alpha,\beta,P_0,\cdots,P_l)$.

$$
\begin{aligned}
&(0\ \beta\ \alpha\ 0)[(I_2\otimes P_0)(I_2\otimes F_0)]=0\\
&\left\{\xi(0\ 1\ 0\ 0)(I_4\otimes P_{s-1})+(0\ \beta\ \alpha\ 0)\sum_{i=0,j=0}^{i+j=s}[(I_2\otimes P_i)(I_{2^{i+1}}\otimes F_{s-i})]\right.\\
&\left.+(0\ 1\ 0\ 0)\sum_{i=1,j=0}^{i+j=s}[(I_4\otimes P_i\Phi_{i-1})(I_{2^{i+1}}\otimes F_{s-i})]\right\}\mathrm{T}_N(2,s+1)=0,\\
&\qquad s=1,\cdots,l+1.\\
&\left\{(0\ \beta\ \alpha\ 0)\sum_{i=0,j=0}^{i+j=s}[(I_2\otimes P_i)(I_{2^{i+1}}\otimes F_{s-i})]\right.\\
&\left.+(0\ 1\ 0\ 0)\sum_{i=1,j=0}^{i+j=s}[(I_4\otimes P_i\Phi_{i-1})(I_{2^{i+1}}\otimes F_{s-i})]\right\}\mathrm{T}_N(2,s+1)=0,\\
&\qquad s=l+2,\cdots,l+k.
\end{aligned}
\tag{12.2.3}
$$

注 这种方法的优点在于:

(1) 它可以通过计算机数值解出;

(2) 这种方法可以很容易推广到高维情形.

12.3 控制系统的非正则线性化

考虑仿射非线性系统

$$\dot{x} = f(x) + \sum_{i=1}^{m} g_i(x)u_i, \quad f(0) = 0, \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m. \tag{12.3.1}$$

假设向量场$f(x)$, $g_i(x)$ 和控制等都是解析的(C^ω). 一般的线性化问题定义如下.

定义 12.3.1 系统(12.3.1) 在原点是非正则状态反馈可线性化的, 如果存在一个反馈控制

$$u = \alpha(x) + \beta(x)v, \tag{12.3.2}$$

其中$\beta(x)$ 是$m \times k$ 矩阵, 和一个微分同胚$z = \xi(x)$, 使得在坐标卡z 下闭环系统可以表示成一个完全能控的线性系统. 如果$\beta(x)$ 是一个非异方阵, 称之为正则状态反馈线性化. 当$k = 1$ 时, 称之为单输入线性化.

Heymann 引理[62] 说, 一个完全能控的线性系统, 一定存在线性非正则反馈

$$u = Hv, \quad v \in \mathbb{R}$$

使反馈系统成为完全能控的单输入系统. 下面的引理简化了非正则状态反馈线性化的研究, 是Heymann 引理的一个直接结果.

引理 12.3.2 系统(12.3.1) 是非线性状态反馈可线性化的, 当且仅当它是单输入可线性化的, 即通过控制(12.3.2) 可线性化, 其中$\beta(x)$ 是$m \times 1$ 的向量.

下面的引理在后面很有用, 它的证明也很直接.

引理 12.3.3 [95] 设$A = J_f(0)$ 是f 在原点的Jacobi 矩阵, $B = g(0)$. 如果系统(12.3.1) 是可线性化的, 那么(A, B) 是完全能控的.

正则型已经被许多学者用来研究非线性系统的许多控制问题, 例如文献[75], [20], [43]. 我们首先介绍一些概念[56]: 设H_n^k 是$\mathbb{R}^n$ 上的一组k 阶齐次多项式向量场, 那么显然有下面的事实:

(1) H_n^k 是$\mathbb{R}$ 上的线性向量空间;

(2) 设$Ax \in H_n^1$ 是一个给定的向量场, 且A 是一个$n \times n$ 常数矩阵. 那么李导数$\mathrm{ad}_{Ax} : H_n^k \to H_n^k$ 是一个线性映射.

下面的正则型表现[3] 和它在线性化方面的应用[43] 是我们的方法的出发点.

我们需要"共振" 与"非共振" 矩阵的概念. 对于一个矩阵A, 设$\sigma(A) = \lambda = (\lambda_1, \cdots, \lambda_n)$ 是它的特征值. A 是一个共振矩阵, 如果存在$m = (m_1 \cdots, m_n) \in \mathbb{Z}_+^n$

和$|m| \geq 2$, 即$m_i \geq 0$ 和$\sum_{i=1}^{n} m_i \geq 2$, 使得对于某个$s$, $\lambda_s = \langle m, \lambda \rangle$. 如果$A$ 不是一个共振矩阵, 则称其为非共振矩阵.

定理 12.3.4 (Poincaré 定理[3])　*考虑C^ω 动态系统*

$$\dot{x} = Ax + f_2(x) + f_3(x) + \cdots, \quad x \in \mathbb{R}^n, \tag{12.3.3}$$

其中$f_i(x)$, $i \geq 2$ 是i 阶齐次多项式向量场. 如果A 是非共振的, 那么存在一个坐标变换

$$x = y + h(y), \tag{12.3.4}$$

其中$h(y)$ 是可能的无限个齐次多项式向量$h_m(y)$, $m \geq 2$ 的和, 即$h(y) = h_2(y) + h_3(y) + \cdots$, 使得系统(12.3.3) 可以表示成$\dot{y} = Ay$.

这个定理已经推广到单输入控制系统[74].

下面的命题对于非共振给出了一个充分条件.

命题 12.3.5 [43]　*设$\lambda = (\lambda_1, \cdots, \lambda_n)$ 是一个给定Hurwitz 矩阵A 的特征值. A 是非共振的, 如果*

$$\max\{|\operatorname{Re}(\lambda_i)| \mid \lambda_i \in \sigma(A)\} \leq 2\min\{|\operatorname{Re}(\lambda_i)| \mid \lambda_i \in \sigma(A)\}. \tag{12.3.5}$$

12.4　单输入线性化

我们首先考虑非正则反馈线性系统的规范型.

称一个常数向量$b = (b_1, \cdots, b_n)^{\mathrm{T}} \in \mathbb{R}^n$ 是分量非零的, 如果$b_i \neq 0$, $\forall i$.

命题 12.4.1　*一个线性控制系统*

$$\dot{x} = Ax + \sum_{i=1}^{m} b_i u_i := Ax + Bu, \quad x \in \mathbb{R}^n, \ u \in \mathbb{R}^m \tag{12.4.1}$$

是完全能控的, 当且仅当存在两个矩阵F, G 使得闭环系统

$$\dot{x} = (A + BF)x + BGv$$

通过线性坐标变换可以转化为如下形式

$$\dot{z} = Az + bv := \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix} z + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} v, \tag{12.4.2}$$

其中$d_i, i=1,\cdots,n$ 互不相同, b 的分量非零.

证明很简单. 关键是这样一个系统的能控性矩阵C

$$\det(C)=\prod_{i=1}^{n}b_i\prod_{i<j}(d_j-d_i)\neq 0. \tag{12.4.3}$$

因为它实质上是一个Vandermonde 矩阵. 由上面的命题我们称(12.4.2) 是非正则反馈单输入A-对角(RSIFAD) 规范形. 进一步, 我们给出下面的假设:

A1. A 是一个对角矩阵, 对角元d_i 互不相同, 并且A 是非共振的.

引理 12.4.2 假设矩阵A 满足A1, g 是一个k 阶齐次向量场, $k\geq 2$. 则存在一个k 阶齐次向量场η 使得

$$\mathrm{ad}_{Ax}\,\eta=g. \tag{12.4.4}$$

证明 对于给定的η, 设$f=ad_{Ax}\eta$, 那么直接计算有f 的第i 个分量f_i 仅依赖于η 的第i 个分量η_i. 现在设$x_1^{r_1}\cdots x_n^{r_n}$ 是η_i 的一项, 直接计算有

$$\mathrm{ad}_{Ax}\,\eta=\left[\begin{pmatrix}d_1x_1\\ \vdots\\ d_ix_i\\ \vdots\\ d_nx_n\end{pmatrix},\begin{pmatrix}\times\\ \vdots\\ x_1^{r_1}\cdot x_n^{r_n}\\ \vdots\\ \times\end{pmatrix}\right]=\begin{bmatrix}\times\\ \vdots\\ \mu_ix_1^{r_1}\cdot x_n^{r_n}\\ \vdots\\ \times\end{bmatrix}, \tag{12.4.5}$$

其中

$$\mu_i=d_1r_1+\cdots+d_nr_n-d_i, \tag{12.4.6}$$

由于A 是非共振的, 且$\mu_i\neq 0$, 对于g_i 的每一项$x_1^{r_1}\cdots x_n^{r_n}$, 我们可以构造$\eta_i$ 的相应的项$\frac{1}{\mu_i}x_1^{r_1}\cdots x_n^{r_n}$, 使得$\mathrm{ad}_{Ax}\,\eta=g$. □

由于假设所有的向量场和函数都是解析的, 那么所有的函数和它们的导数都有收敛的Taylor 级数展开.

注意到如果A 满足A1, 那么对于向量场$g=g_kx^k+g_{k+1}x^{k+1}+\cdots\in O(\|x\|^k)$, 对它的每一项应用引理12.4.2, 我们可以找到一个向量场$\eta\in O(\|x\|^k)$ 使得$\mathrm{ad}_{Ax}\,\eta=g$.

现在让我们回到线性化. 我们考虑下面的系统:

$$\dot{x}=Ax+\xi(x)+\sum_{i=1}^{m}g_i(x)u_i, \tag{12.4.7}$$

其中A 满足A1, 且$\xi(x) = O(\|x\|^2)$. 一个直接结论是

命题 12.4.3　考虑系统(12.4.7). 它是非正则反馈可线性化的, 如果
(1) $\xi(x) \in \mathrm{Span}\{g_1, \cdots, g_m\}$;
(2) 存在一个分量非零的常数向量b, 使得

$$b \in \mathrm{Span}\{g_1, \cdots, g_m\}.$$

当命题12.4.3 中有一个条件不成立时, 我们可以用正则型来进一步研究这个问题.

根据引理12.4.2, 我们总是可以找到一个向量场$\eta(x)$ 使得

$$\mathrm{ad}_{Ax}\,\eta(x) = \xi(x). \tag{12.4.8}$$

现在我们定义一个局部微分同胚$z_1 = x - \eta(x)$, 于是在坐标卡z_1 下系统(12.4.7) 可以表示成

$$\dot{z}_1 = Az_1 - J_0(x)\xi(x) + \sum_{i=1}^{m} g_i^1(x)u_i, \tag{12.4.9}$$

其中$J_0(x)$ 是$\eta(x)$ 的Jacobi 矩阵, 且$g_i^1(x) = (I - J_0(x))g_i(x)$.

为了记号的简便, 我们记$x := z_0$, $\xi(x) := \xi_0(x)$, $\eta(x) := \eta_0(x)$, $g_i(x) := g_i^0(x)$. 于是我们可以继续前面的步骤来递推定义新的坐标

$$\mathrm{ad}_{Ax}(\eta_k) = \xi_k, \quad z_{k+1} = z_k - \eta_k(x), \quad k \geq 0$$

和新的向量场

$$g_i^{k+1}(x) = (I - J_k(x))\, g_i^k(x), \quad 1 \leq i \leq m, \quad k \geq 0,$$

其中$J_k(x)$ 是$\eta_k(x)$ 的Jacobi 矩阵. 这样在z_k 下, 系统可表示成

$$\dot{z}_k = Az_k + \xi_k(x) + \sum_{i=1}^{m} g_i^k(x)u_i, \quad k \geq 1. \tag{12.4.10}$$

作为上面讨论的结果, 我们有

推论 12.4.4　系统(12.4.7) 是非正则状态反馈可线性化的, 如果存在$k \geq 0$ 使得(12.4.10) 满足命题12.4.3 中的条件(1) 和(2).

根据递推算法容易看出

$$deg(\xi_i) = c_{i+1} + 1, \quad i = 0, 1, \cdots,$$

其中$\{c_i\}$ 是Fibonacci 序列, 即$(c_1, c_2, \cdots) = (1, 1, 2, 3, 5, 8, \cdots)$. 因此当$k \to \infty$ 时$\xi_k(x) \to 0$, 因为已经假设它是收敛的. 于是我们有

推论 12.4.5 系统(12.4.7) 是非正则状态反馈可线性化的, 如果存在一个分量非零的常数向量b 使得

$$b \in \text{Span}\left\{\prod_{i=0}^{\infty}(I - J_i(x))g_j(x),\ j = 1, \cdots, m\right\}.$$

12.5 非正则反馈线性化算法

本节我们将首先给出实现Poincare 坐标变换(12.3.4) 的公式, 然后给出(近似)线性化的充要条件.

首先, 对$f(x)$ 利用半张量积形式的Taylor 级数展开, 我们可以将系统(12.3.3) 表示成

$$\dot{x} = Ax + F_2x^2 + F_3x^3 + \cdots, \tag{12.5.1}$$

其中F_k 是$n \times n^k$ 常数矩阵, x^k 如前所述.

接着, 假设

$$\text{ad}_{Ax}\,\eta_k = F_kx^k.$$

利用引理12.4.2, 我们可以得到

$$\eta_k = (\Gamma_k^n \odot F_k)\,x^k, \quad x \in \mathbb{R}^n. \tag{12.5.2}$$

这里$\odot$ 是矩阵的Hadamard 积(参见第2 章). Γ_k^n 可由(12.4.6) 构造如下

$$(\Gamma_k^n)_{ij} = \frac{1}{\left(\sum\limits_{s=1}^{n}\alpha_s^j\lambda_s\right) - \lambda_i}, \quad i = 1, \cdots, n;\ j = 1, \cdots, n^k. \tag{12.5.3}$$

其中$\alpha_1^j, \cdots, \alpha_n^j$ 分别是x^k 的第j 个分量$x_1, \cdots, x_n$ 的幂.

现在我们将介绍我们的主要结果.

定理 12.5.1 假设A 满足A1, 则系统(12.5.1) 可以通过下面的坐标变换

$$z = x - \sum_{i=2}^{\infty} E_ix^i \tag{12.5.4}$$

转换为线性形式

$$\dot{z} = Az, \tag{12.5.5}$$

其中E_i 由下式递推决定(Φ_i 如第5 章(5.5.5) 中所示)

$$\begin{aligned} E_2 &= \Gamma_2 \odot F_2, \\ E_s &= \Gamma_s \odot \left(F_s - \sum_{i=2}^{s-1} E_i\Phi_{i-1}(I_{n^{i-1}} \otimes F_{s+1-i})\right),\ s \geq 3. \end{aligned} \tag{12.5.6}$$

证明　对系统应用(12.5.4), 我们有

$$
\begin{aligned}
\dot{z} &= \left(Ax+\sum_{i=2}^{\infty}F_ix^i\right)-\sum_{i=2}^{\infty}\frac{\partial E_ix^i}{\partial x}\left(Ax+\sum_{i=2}^{\infty}F_ix^i\right)\\
&= Az+\sum_{i=2}^{\infty}F_ix^i+A\sum_{i=2}^{\infty}E_ix^i-\sum_{i=2}^{\infty}\frac{\partial E_ix^i}{\partial x}Ax\\
&\quad -\left(\sum_{i=2}^{\infty}\frac{\partial E_ix^i}{\partial x}\right)\left(\sum_{j=2}^{\infty}F_jx^j\right)\\
&= Az-\sum_{i=2}^{\infty}\mathrm{ad}_{Ax}(E_ix^i)+F_2x^2+\sum_{s=3}^{\infty}\Big(F_sx^s-\\
&\quad \sum_{i=2}^{s-1}\frac{\partial E_ix^i}{\partial x}F_{s+1-i}x^{s+1-i}\Big)\\
&:= Az-\sum_{i=2}^{\infty}\mathrm{ad}_{Ax}(E_ix^i)+\sum_{s=2}^{\infty}L_s,
\end{aligned}
\tag{12.5.7}
$$

其中

$$
\begin{aligned}
L_2 &= F_2x^2\\
L_s &= F_sx^s-\sum_{i=2}^{s-1}\frac{\partial E_ix^i}{\partial x}\ltimes F_{s+1-i}x^{s+1-i}\\
&= \left(F_s-\sum_{i=2}^{s-1}E_i\Phi_{i-1}(I_{n^{i-1}}\otimes F_{s+1-i})\right)x^s,\ s\geq 3.
\end{aligned}
\tag{12.5.8}
$$

在假设*A1* 下我们可以令

$$E_sx^s=\mathrm{ad}_{Ax}^{-1}(L_s),\quad s=2,3,\cdots.$$

于是(12.5.1) 变成(12.5.5).　□

这种Taylor 级数展开的优势是不需要在变换坐标z_i, $i=1,2,3,\cdots$ 下递推计算系统的中间形式.

现在我们考虑系统(12.3.1) 的线性化. 记$A=\frac{\partial f}{\partial x}|_0$, $B=g(0)$, 并且假设(A,B) 是完全能控对. 那么我们可以找到反馈K 和线性坐标变换T 使得$\tilde{A}=T^{-1}(A+BK)T$ 满足假设*A1*. 为了简单起见, 我们称上面的变换为非共振型变换(简记为NR 型变换).

利用前面的记号和算法立即得到下面的结论.

定理 12.5.2 系统(12.3.1) 是单输入可线性化的, 当且仅当存在一个NR 型变换和分量非零的常数向量b 使得

$$b \in \operatorname{Span}\left\{ \left(I - \sum_{i=2}^{\infty} E_i \Phi_{i-1} x^{i-1} \right) g_j \middle| j = 1, \cdots, m \right\}. \tag{12.5.9}$$

下面, 我们考虑近似线性化.

定义 12.5.3 称系统(12.3.1) 是k 阶非正则状态反馈近似可线性化的, 如果我们可以找到一个状态反馈和一个局部坐标卡z, 使得在z 下闭环系统可以表示成

$$\dot{z} = Az + O(\|z\|^{k+1}) + (b + O(\|z\|^k))v, \tag{12.5.10}$$

其中(A, b) 是完全能控对.

对于近似可线性化, 我们可以放松非共振的约束.

定义 12.5.4 设$\lambda = (\lambda_1, \cdots, \lambda_n)$ 是矩阵A 的特征值, A 是k 阶共振矩阵, 如果存在$m = (m_1 \cdots, m_n) \in \mathbb{Z}_+^n$ 和$2 \le |m| \le k$, 使得对于某个s, $\lambda_s = \langle m, \lambda \rangle$.

从表达式(12.5.3) 可以看出下面的Poincare 定理的推论是正确的.

推论 12.5.5 考虑C^ω 动态系统(12.3.3). 如果A 是k 阶非共振的, 则存在坐标变换(12.3.4), 使得系统(12.3.3) 可以表示成

$$\dot{z} = Az + O(\|z\|^{k+1}). \tag{12.5.11}$$

如果我们只考虑系统(12.5.1) 在$u_i = 0$ 时的k 阶近似线性化, 那么我们只需调整(12.5.4) 为

$$z = x - \sum_{i=2}^{k} E_i x^i. \tag{12.5.12}$$

在这种情况下, (12.5.6) 仍然有效(这里$s \le k$), 并且(12.5.5) 将变成

$$\dot{z} = Az + O(\|x\|^{k+1}). \tag{12.5.13}$$

我们称一个变换是非共振 k 型变换 (简记NR-k 型变换), 如果它和NR 型变换几乎相同, 除了用k 阶非共振取代非共振条件外.

定理 12.5.6 系统(12.3.1) 是k 阶单输入状态反馈近似可线性化的, 当且仅当存在一个NR-k 型变换和一个分量非零的常数向量b 使得

$$b \in \operatorname{Span}\left\{ \left(I - \sum_{i=2}^{k} E_i \Phi_{i-1} x^{i-1} \right) g_j \middle| \forall j \right\} + O(\|x\|^k). \tag{12.5.14}$$

我们用下面的例子来说明线性化的过程.

例 12.5.7 找到如下控制系统的一个4 阶近似状态反馈线性化

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{bmatrix} = \begin{bmatrix}-4\sin x_1 - \frac{2}{3}x_1^3 + 5x_2^2 + 6x_2^3\\ -5x_2 - 3x_3^2\\ -6x_3\end{bmatrix} + \begin{bmatrix}0\\ 6(1+x_3)\\ 7\end{bmatrix}u_1 + \begin{bmatrix}1\\ 0\\ 0\end{bmatrix}u_2. \tag{12.5.15}$$

利用Taylor 级数展开, 我们可以将(12.5.15) 表示成

$$\dot{x} = \begin{bmatrix}-4 & 0 & 0\\ 0 & -5 & 0\\ 0 & 0 & -6\end{bmatrix}x + \begin{bmatrix}5x_2^2\\ -3x_3^2\\ 0\end{bmatrix} + \begin{bmatrix}6x_2^3\\ 0\\ 0\end{bmatrix} + O(\|x\|^5) + \begin{bmatrix}0\\ 6+6x_3\\ 7\end{bmatrix}u_1 + \begin{bmatrix}1\\ 0\\ 0\end{bmatrix}u_2. \tag{12.5.16}$$

容易算出

$$L_2 = (5x_2^2,\ -3x_3^2,\ 0)^{\mathrm{T}},$$

$$E_2x^2 = ad_{Ax}^{-1}(L_2) = (-\tfrac{5}{6}x_2^2,\ \tfrac{3}{7}x_3^2,\ 0)^{\mathrm{T}};$$

$$L_3 = (6x_2^3 - 5x_2x_3^2,\ 0,\ 0)^{\mathrm{T}},$$

$$E_3x^3 = ad_{Ax}^{-1}(L_3) = \left(-\tfrac{6}{11}x_2^3 + \tfrac{5}{13}x_2x_3^2,\ 0,\ 0\right)^{\mathrm{T}}.$$

因此我们得到希望的坐标变换如下

$$z = x - \begin{bmatrix}-\frac{5}{6}x_2^2 - \frac{6}{11}x_2^3 + \frac{5}{13}x_2x_3^2\\ \frac{3}{7}x_3^2\\ 0\end{bmatrix}, \tag{12.5.17}$$

在这个坐标变换下系统(12.5.15) 可以表示成

$$\dot{z} = \begin{bmatrix}-4 & 0 & 0\\ 0 & -5 & 0\\ 0 & 0 & -6\end{bmatrix}z + O(\|x\|^5) + \begin{bmatrix}h(x) & 1\\ 6 & 0\\ 7 & 0\end{bmatrix}\begin{bmatrix}u_1\\ u_2\end{bmatrix}, \tag{12.5.18}$$

其中$h(x)=(6+6x_3)(\frac{5}{3}x_2+\frac{18}{11}x_2^2-\frac{5}{13}x_3^2)-\frac{70}{13}x_2x_3$. 因为

$$\begin{bmatrix}1\\6\\7\end{bmatrix}=\begin{bmatrix}h(x)\\6\\7\end{bmatrix}\times 1+\begin{bmatrix}1\\0\\0\end{bmatrix}\times(-h(x)+1),$$

定理12.5.6 保证系统是4 阶状态反馈近似可线性化的.

选择状态反馈

$$\begin{bmatrix}u_1\\u_2\end{bmatrix}=\begin{bmatrix}1\\-h(x)+1\end{bmatrix}v. \tag{12.5.19}$$

并将它代入到(12.5.18) 中, 我们得到

$$z=\begin{bmatrix}-4&0&0\\0&-5&0\\0&0&-6\end{bmatrix}z+O(||x||^5)+\begin{bmatrix}1\\6\\7\end{bmatrix}v. \tag{12.5.20}$$

这就是我们要找的系统(12.5.15) 的4 阶状态反馈近似线性化. □

12.6 注释与参考

需要了解关于Carleman 线性化更多信息和后续进展请参考文献[93] 和[12].

用Carleman 线性化技术讨论一次积分有许多讨论, 特别是对一些特殊系统, 一次积分有其物理意义. 这方面的工作, 例如, 在文献[93] 中讨论了Lorenz 系统的这种一次积分, 在文献[12] 中讨论了Lotka-Volterra 系统的一次积分.

非正则反馈线性化比正则反馈线性化更强有力, 有关文献如[95], [51]. 但远不像正则反馈线性化, 非正则反馈线性化至今尚未有较系统的结果. 因此, 非正则状态反馈线性化仍是一个富有挑战性的问题.

本章内容是基于以下几篇文章: Carleman 线性化来源于文献[26], 非正则反馈线性化的结果主要来自文献[27].

参考文献

[1] Abraham RA, Marsden JE. *Foundations of Mechanics.* 2nd Ed. Benjamin/Cummings Pub. Com. Inc., 1978

[2] Aeyels D. Stabilization of a class of nonlinear systems by a smooth feedback control. *Sys. Contr. Lett.*, Vol. 5: 189~294, 1985

[3] Arnold VI. *Geometrical Methods on the Theory of Ordinary Differential Equations.* Mew York, Springer-Verlag, 177 ~ 188, 1983

[4] Babuska R. An overview of fuzzy modeling and model-based fuzzy control, in *Fuzzy logic Control, Advances in Applications*, H.B. Verbruggen (ed.), World Scientific Publishing, Singapore, 1999

[5] Bates DM, DG Watts. Relative curvature measure of nonlinearity. *J.R. Statist.*, Soc. B 42: 1~25, 1980

[6] Bates DM, DG Watts. Parameter transformations for improved approximate confidence regions in nonlinear least squares. *Ann. Statist.*, 9: 1152~1167, 1981

[7] Bates L, J Sniatycki. Nonholonomic reduction. Reports on Math. Phys., 32: 99~115, 1993

[8] Bellman, R., *Introduction to Matrix Analysis*, 2nd Ed. McGraw-Hill Book Comp. 2nd Ed. 1970

[9] Bluman GW, S Kumei. *Symmetries and Differential Equations.* Springer-Verlag, 1989

[10] Boothby WM. *An Introduction to Differentiable Manifolds and Riemannian Geometry.* 2nd Ed. Academic Press, 1986

[11] Byrnes CI, Isidori A. Asymptotic stabilization of minimum-phase nonlinear systems. *IEEE Trans. Aut. Contr.*, Vol. 36: 1122~1137, 1991

[12] Cairo L, MR Feix. Families of invariants of the motion for Lotka-Volterra 1st family. *J. Math. Phys.*, Vol. 33, No. 7, 2240~2455, 1992

[13] Cairo L, MR Feix, J Llibre. Darboux method and rearch of invariats for the Lotka-Volterra and complex quadratic systems. *J. Math. Phys.*, Vol. 40: No. 4, 2074~2091, 1999

[14] Carr J. *Applications of Center Manifold Theory.* Springer-Verlag, 1981

[15] Cheng D, W Xue, J Huang. On general Hamiltonian systems. *Proc. ICARCV 98*, Singapore, 185~189, 1998

[16] Cheng D, Z Xi, Q Lu, S Mei. Geometric structure of generalized controlled Hamiltonian systems and its application. *Science in China*, Series E, Vol. 43, No. 4, 365~379, 2000

[17] Cheng D. Semi-tensor product of matrices and its application to Morgan's problem. Science in China, Series F, Vol. 44, No. 3, 195~212, 2001

[18] Cheng D. On Lyapunov Mapping and its Applications. *Commu. on Inform. Sys.*, Vol. 1, No. 3, 255~272, 2001

[19] Cheng D, C Martin. Stabilization of nonlinear systems via designed center manifold. *IEEE Trans. Aut. Contr.*, Vol. 46, No. 3, 1372~1383, 2001

[20] Cheng D. *Matrix and Polynomial Approach to Dynamic Control Systems.*Beijing, Science Press, 2002

[21] Cheng D, L Zhang. On semi-tensor product of matrices and its applications. *Acta Math Appl Sinica*, Vol. 19, No. 2, 219~228, 2003

[22] Cheng D. Stabilization via polynomial Lyapunov function, in *New Trends in Nonlinear Dynamics and Control*, Wei Kang, Mingqing Xiao, Carlos Borges (Eds.) Berlin, Springer, 161~174, 2003

[23] Cheng D, L Zhang. Generalized normal form and stabilization of nonlinear systems. *Int. J. Control*, Vol. 76, No. 2, 116~128, 2003

[24] Cheng D, L Guo, J. Huan. On Quadratic Lyapunov Functions. *IEEE Trans. Aut. Contr.*, Vol. 48, No. 5, 885~890, 2003

[25] Cheng D, J Ma. Calculation of Stability Region. *Proc.* 42*nd IEEE CDC*'03 , Maui, 5615~5620, 2003

[26] Cheng D, Y Dong. Semi-tensor Product of Matrices and Its Some Applications to Physics. *Methods and Applications of Analysis*, Vol. 10, No. 4, 565~588, 2003

[27] Cheng D, X Hu, Y Wang. Non-regular feedback linearization of nonliner systems via a normal form algorithm. *Automatica*, Vol. 40, No. 3, 439~447, 2004

[28] Cheng D, Y Dong. Semi-tensor product of matrices and its some applications to physics. *Matheds Appl. of Analysis*, 2004, in Press

[29] Cheng D, J Ma, Q Lu, S Mei. Quadratic form of stable sub-manifold for power systems. *Int. J. Rob. Nonl. Cont.*, 2004, in Press

[30] Cheng D, Z Xi, G Feng. Stabilization of general nonlinear control systems via center manifold and approximation techniques. *JDCS*, Vol. 10, No.3, 315~327, 2004

[31] Cheng D. On Logic-based Intelligent Systems. *Proc. ICCA 2005*, Budapest, 71~76, 2005

[32] Cheng D, H Qi. Matrix Expression of Logic and Fuzzy Control. *Proc.* 44*th IEEE CDC*, Seville, 3273~3278, 2005

[33] Cheng D. Some applications of semitensor product of matrices in algebra. *Comp. Math. Appl.*, to appear

[34] Cheng D, G Yang, Z Xi. Nonlinear systems possissing linear symmetry. *Int. J. Robust, Nonlin. Contr.*, to appear

[35] Chiang HD, M Hirsch, F Wu. Stability regions of nonlinear autonomous dynamical systems. *IEEE Trans. Automat. Contr.*, Vol. 33, No. 1, 16~27, 1988

[36] Chiang HD, FF Wu. Foundations of the potential energy boundary surface method

for power system transient stability analysis. *IEEE Trans. Circ. Sys.*, Vol. 35, No. 6, 712～728, 1988

[37] Chiang HD, Chia-Chi Chu. Theoretical foundation of the BCU method for direct stability analysis of network-reduction power system models with small transfer conductances. *IEEE Trans. Circ. Sys.*, Vol. 42, No. 5, 252～265, 1995

[38] Chiang HD, Chia-Chi Chu, G Cauley. Direct stability abalysis of electric power systems using energy functions: theory, applications, and perspective. *Proceedings of the IEEE*, Vol. 83, No. 11, 1497～1529, 1995

[39] Conway JB. *A Course in Functional Analysis.* Springer-Verlag, New York, 1985

[40] Deif AS. *Advanced Matrix Theory for Scientists and Engineers.* New York: John Wiley & Sons, 1982

[41] Descusse J, CH Morg. Decoupling with dynamic comp ensation for statrong invertible affine nonlinear systems. *Int. J. Contr.*, Vol. 43, 1385～1398, 1985

[42] Descusse J, JF Lafay, M Malabre. Solution to Morgan's problem. *IEEE Trans. Aut. Contr.*, Vol. 33, No. 8, 732～739, 1988

[43] Devanathan R. Linearization condition through state feedback. *IEEE Trans. Aut. Contr.*, Vol. 46, No. 8, 1257～1260, 2001

[44] Di Benedetto MD, A Glumineau, CH Moog. The nonlnear interactor and its application to imput-output decoupling. *IEEE Trans. Aut. Contr.*, Vol. 39, No. 9, 1240～1250, 1994

[45] Dong Y, D Cheng, H Qin. Feedback stabilization via designed planar center manifold. *Int. J. Robust Nonlin. Contr.*, Vol.14, 1～14, 2004

[46] Eisenhart LP. *Riemannian Geometry.* Princeton University Press, New Jersey, 1997

[47] Falb PL, WA Wolovich. Decoupling and synthesis of multivariable control systems. *IEEE Trans. Aut. Contr.* Vol. 12, No. 5, 651～669, 1967

[48] Foster J, JD Nightngale. *A Short Course in General Relativity.* Springer-Verlag, 1995

[49] Freund E. The structure of decoupled nonlinear systems. *Int. J. Contr.*, Vol. 21, 443～450, 1975

[50] Gardner RB, WF Shadwick. Symmetry and the implementation of feedback linearization. *Sys. Contr. Lett.*, Vol. 15, 25～33, 1990

[51] Ge SS, Z Sun, TH Lee. Nonregular feedback linearization for a class of secondorder nonlinear systems. *Automatica*, Vol. 37, 1819～1824, 2001

[52] Gerla G. *Fuzzy Logic, Mathematical Tools for A pproximate Reasoning.* Kluwer, 2001

[53] Glumineau A, CH Moog. Nonlinear Morgan's problem: Case of $p+1$ inputs and p outputs. *IEEE Trans. Aut. Contr.*, Vol. 37, No. 7, 1067～1072, 1992

[54] Greub W. *Multilinear Algebra.* 2nd Ed. Springer-Verlag, New York, 1978

[55] Grizzle JW, SI Marcus. The structure of nonlinear control systems possessing symmetries. *IEEE Trans. Aut. Contr.* Vol. 30, No. 3, 248～258, 1985

[56] Guckenheimer J, P Holmes. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields.*Berlin: Springer, 1983

[57] Hahn W. Stability of Motion. Springer-Verlag, 1967

[58] Hamilton AG. *Logic for Mathematicians.* Revised Ed. Cambridge: Cambridge University Press, 1988

[59] Hartshorne R. *Algebraic Geometry.* Springer-Verlag, New York, 1977

[60] Herera HAN, JF Lafay. New results about Morgan's problem, Vol. 38, No. 12, 1834~1838, 1993

[61] Hermes H. Resonance and feedback stabilization. *Proc. IFAC Nonlinear Control Systems Design Symposium,* (NOLCOS'95), Tahoe, 47~52, 1997

[62] Heymann M. Pole assignment in multi-input linear systems. *IEEE Trans. Aut. Contr.*, Vol. 13, No. 6, 748~749, 1968

[63] Horn R, C Johnson. *Topics in Matrix Analysis.* Cambridge:Cambridge University Press, 1991

[64] Huang J. *Nonlinear Output Regulation, Theory and Applications.* Siam, Philadelphia, 2004

[65] Huang L. *Linear Algebra in Systems and Control Theory.* Science Press (in Chinese), 1984

[66] Humphreys JE. *Introduction to Lie Algebras and Representation Theory.* Springer-Verlag, 1972

[67] Hungerford TW. *Algebra.* Springer-Verlag, 1974

[68] Isidori A. *Non-linear Control Systems.* 3rd Ed. Springer-Verlag, 1995

[69] Jurdjevic V. Optimal control problems on Lie groups: Crossroads between geometry and methanics, in B.Jakubczyk, W.Respondek (Eds.). *Geometry of Feedback and Optimal Control,* Marcel Dekker, New York, 257~303, 1997

[70] Khalil HK. *Nonlinear Systems.* Macmillan Publishing Company, New York, 1992

[71] Khalil HK. *Nonlinear Systems.* 3rd Ed. Prentice Hall, 1996

[72] Knobloch HW, A Isidori, D Flockerzi. *Topics in Control Theory.* Basel: Birkhäuser Verlag, 1993

[73] Koon WS, JE Marsden. Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction. *SIAM J. Control and Optim.*, 1996

[74] Krener AJ, S Karahan, M Hubbard. Approximate normal forms of nonlinear systems. *Proc. 27th Conf. Dec. Contr.*, 1223~1229, 1988

[75] Krener AJ, W Kang. Extended normal forms of quadratic systems. *Proc, 29th Conf. Dec. Contr.*, 2091~2095, 1990

[76] Kuipers JB. Quaternions and Ratation Sequences. Princeton: Princeton Univ. Press, 1999

[77] Lang S. *Algebra.* Revised 3rd Ed. New York: Springer-Verlag, 2002

[78] Myerson RB. *Game Theory, Analysis of Conflict.* Harvard Univ. Press, London, 1991

[79] Liu Z, Y Liu. *Fuzzy Logic and Neural Network.* BUAA Press, Beijing, 1996 (in Chinese)

[80] Magnus JR, H Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Revised Ed. John Wiley & Sons, 1999

[81] Marsden JE, TS Ratiu. *Introduction to Mechanics and Symmetry.* Springer-Verlag, New York, 1994

[82] Morgan BS. The synthesis of linear multivariable systems by state feedback. *J.A.C.C.*, 1964, Vol. 64, 468~472, 1964

[83] Olver PJ. *Applications of Lie Groups to Differential Equations.* Springer-Verlag, 1986

[84] Olver PJ. *Equivalence, Invariants, and Symmetry.* Cambridge Univ. Press, 1995

[85] Ooba T, Y Funahashi. Two conditions concerning common quadratic Lyapunov functions for linear systems. *IEEE Trans. Aut. Contr.*, Vol. 42, No. 5: 719~721, 1997

[86] Passino KM, S Yurkovich. *Fuzzy Control.* Addison Wesley Longman, 1998

[87] Rade L, B Westergren. *Mathematics Handbook for Science and Engineering.* 4th Ed. Sweden: Studentlitteratur, 1998

[88] Respondek W, IA Tall. How many symmetries does admit a nonlinear single-input control system around an equilibrium. *Proc. 40th IEEE CDC*, Orlando, 1795~1800, 2001

[89] Respondek W, IA Tall. Nonlinearizable single-input control systems do not admit stationary symmetries. *Sys. Contr. Lett.*, Vol. 46, 1~16, 2002

[90] Sachs RK, H Wu. *General Relativity for Mathematicians.* Springer-Verlag, 1977

[91] Saha S, AA Fouad, WH Kliemann, V Vittal. Stability boundary approximatin of a power system using the real normal form of vector fields. *IEEE Trans. Power Sys.*, Vol.12, No. 2, 797~802, 1997

[92] Sokolnikoff IS. *Tensor Analysis, Theory and Applications to Geometry and Mechanics of Continua.* 2nd Ed. John Wiley& Sons, Inc., 1964

[93] Steeb WM, F Wilhelm. Non-linear autonomous systems of differential equations and Carleman linearization procedure. *J. Math. Anal. Appl.*, Vol. 77, 601~ 611, 1980

[94] Suda M, K Unahashi. Decoupling of nonsquare systems-a necessary and sufficient condition in terms of infinite zeros. *Prec. 9th IFAC World Conference*, Vol. 8, 88~93, 1984

[95] Sun Z, X Xia. On nonregular feedback linearization. *Automatica*, Vol. 33, No. 7, 1739~1744, 1997

[96] Truemper K. *Design of Logic-based Intelligent Systems.* Wiley & Sons, New Jersey, 2004

[97] Tsai CL. *Contributions to the Design and Analysis of Nonlienar Models.* Ph.d. Thesis, Univ. of Minisota, 1983

[98] van der Schaft A. L_2 Gain and Passivity Techniques in Nonlinear Control. New York: Springer-Verlag, 1996

[99] Varadarajan VS. *Lie Groups, Lie Algebras, and Their Representations.* New York: Springer-Verlag,1984

[100] Varaiya PV, FF Wu, R Chen. Direct methods for transient stability analysis of power systems: recent results. *Proc. IEEE*, Vol. 73, No. 12, 1703∼1715, 1985

[101] Venkatasubramanian V, W Ji. Numerical approximation of $(n-1)$-dimensional stable manifolds in large systems such as the power system. *Automatica*, Vol. 33, No. 10, 1877∼1883, 1997

[102] Verbruggen HB, R Babuska (eds). *Fuzzy Logic Control Advances in Applications.* Singapore: World Scientific,1999

[103] Walser H. Symmetry. *Math. Assoc. Amer.*, Washington DC, 2000

[104] Wang L. *A Course in Fuzzy Systems and Control.* Prentice-Hall, 1997

[105] Wang G. *Non-Classical Mathematical Logic and Fuzze Infection.* Beijing: Science Press, 2000

[106] Wang S, Z Jia. *Inequalities in Matrix Theory.* Anhui Education Pub. (in Chinese), 1994

[107] Wang X. *Theory and Applications of Parameter Estimate of Nonlinear Models.* Wuhan: Wuhan University Press, (in Chinese), 2002

[108] Wei B. The second moment of LSE in the nonlinear regression model. *App. Math., J. Chinese Universities*, Vol.1, No. 2, 279∼286, (in Chinese), 1986

[109] Willems JL. *Stability Theory of Dynamical Systems.* John Wiley & Sons Inc., 1970

[110] Wonham WM, AS Morse. Decoupling and pole assignment in linear multivariable systems: A geometric approach. *SIAM J. Control*, Vol. 8, No. 1, 1∼18, 1970

[111] Wonham WM. *Linear Multivariable Control. A Geometric Approach.* Berlin: Springer-Verlag, 1974

[112] Wolovich WA. *Lienar Multivariable Systems.* Springer-Verlag, 1974

[113] Wu W. *On Mechanization of Mathematics.* Jinan: Shandong Educational Press, 1995

[114] Xie X, X Liu, S Zhang. The study of structure and properties of second order nonlinear systems possessing rotation symmetry. *J. Sys. Sci. Math. Sci.* Vol. 13, No. 2, 120∼127, 1993

[115] Xue A, KL Toe, Q Lu, S Mei. Polynomial approximations for the stable and unstable manifolds of hyperbolic equilibrium point using semi-tensor product. *Int. J. Innov. Comp. Inform. Contr.*, Vol. 2, No. 3, 1∼16, 2006

[116] Xue A, FF Wu, Q Lu, S Mei. Power system dynamic security region and its applications. *IEEE Trans. Circ. Sys. I*, to appear

[117] Yokonuma T. *Tensor Spaces and Exterior Algebra.* Providence:American Mathematical Society, 1992

[118] Zaborszky JG, JG Huang. B Zheng, TC Leung. On the phase protraits of a class of large nonlinear dynamic systems such as the power systems. *IEEE Trans. Autom. Contr.*, Vol. 33, No. 1, 4~15, 1988

[119] Zadeh L, J Kacprzyk (ed.). *Fuzzy Logic for the Management of Uncertainty.* John Wiley & Sons, Inc., 1992

[120] Zhang F. *Matrix Theory, Basic Results and Techniques.* New York: Springer-Verlag, 1999

[121] Zhang L, D Cheng. The general structure of cubic matrices, submitted for pub. 2003

[122] Zhang Y. *Multi-matrix Theory.* Chinese statistics Press, Beijing, 1993

[123] Zhao J, S Zhang. On the controllability of nonlinear systems with symmetry. *Sys. Contr. Lett.*, Vol. 18, No. 6, 445~448, 1992

附　　录

本附录给出一些简单函数的计算程序和例子, 是应用MatLab 写成的, 它们为读者学习和应用半张量积提供方便.

A.1　常用函数

1. 计算两个矩阵A, B 的左半张量积:

```
function P=SP(A,B);
Sa=size(A);
Sb=size(B);
n=Sa(2);
p=Sb(1);
if mod(n,p)==0
    z=n/p;
    P=A*kron(B,eye(z));
elseif mod(p,n)==0
    z=p/n;
    P=kron(A,eye(z))*B;
else
    P=0;
end
```

2. 计算n (≥ 2) 个矩阵的左半张量积:

```
function P = SPN(varargin)
n=length(varargin);
if n<2
  error('Not enough input arguments.');
elseif n==2
  P=SP(varargin{1},varargin{2});
else
  P=SP(varargin{1},varargin{2});
  for i=3:n
    P=SP(P,varargin{i});
```

```
  end
end
```

3. 计算换位矩阵$W_{[m,n]}$:

```
function W=WIJ(m,n);
if nargin==1
  n=m;
end
d=m*n;
W=zeros(d);
for k=1:d
  j=mod(k,n);
  if j==0
    j=n;
  end
  i=(k-j)/n+1;
  W((j-1)*m+i,k)=1;
end
```

4. 给出矩阵A 的列展开:

```
function x=VC(A);
x=A(:);
```

5. 给出矩阵A 的行展开:

```
function x=VR(A);
A=A';
x=A(:);
```

6. 给出矩阵列展开的逆(参见定义1.4.7):

```
function A=invVC(x,m)
x=x(:);
l=length(x);
if nargin==1
```

```
  m=ceil(sqrt(l));
  if m^2-l~=0
    x(l+1:m^2)=0;
    l=m^2;
  end
else
  r=mod(l,m);
  if r ~= 0
    x(l+1:l+m-r)=0;
    l=l+m-r;
  end
end
A=reshape(x,m,l/m);
```

7. 给出矩阵行展开的逆(参见定义1.4.7):

```
function A=invVR(x,n)
if nargin==1
  A=invVC(x);
else
  A=invVC(x,n);
end
A=A';
```

8. 计算矩阵A 的块转置$A^{\mathrm{T}(s\times t)}$:

```
function T=BT(A,s,t)
[m,n]=size(A);
if mod(m,s)~=0 | mod(n,t)~=0
  error('Can not perform block transpose.');
end
p=m/s;
q=n/t;
T = zeros(s*q,t*p);
if isa(A,'sym')
  T = sym(T);
```

```
end
for i=1:p
  for j=1:q
    T((j-1)*s+1:j*s,(i-1)*t+1:i*t)=A((i-1)*s+1:i*s,(j-1)*t+1:j*t);
  end
end
```

A.2 算例

1. 计算例1.4.3 中矩阵的块转置:

```
% Example 1.4.3
syms a11 a12 a13 a14 a21 a22 a23 a24 a31 a32 a33 a34 a41 a42 a43 a44
A=[a11 a12 a13 a14;
   a21 a22 a23 a24;
   a31 a32 a33 a34;
   a41 a42 a43 a44];
BT(A,2,2)
BT(A,1,2)
```

运行后的结果是

```
ans =

[ a11, a12, a31, a32]
[ a21, a22, a41, a42]
[ a13, a14, a33, a34]
[ a23, a24, a43, a44]

ans =

[ a11, a12, a21, a22, a31, a32, a41, a42]
[ a13, a14, a23, a24, a33, a34, a43, a44]
```

2. 计算例2.3.6 中矩阵的左半张量积:

```
% Example 2.3.6
```

```
X=[1 2 -1 3];
Y=[2 1]';
SP(X,Y)

X=[2 1];
Y=[1 2 -1 3]';
SP(X,Y)

X=[1 2 -1 2; 0 1 2 3; 3 3 1 1];
Y=[1 2; -1 3];
SP(X,Y)
```

运行后的结果是

```
ans =

     1     7

ans =

     1
     7

ans =

     2     0    -1    10
    -2    -2     6    11
     2     2     9     9
```

3. 计算例3.2.4 中的映射:

```
% Example 3.2.4
A=[1 -1; 2 1; 0 1];
B=[1 -1 0 1; 2 1 1 0];
C=[1 2; 0 1; -1 1];
D=[1 1 1 -1; 1 0 2 1];
```

```
syms z11 z12 z21 z22
Z=[z11 z12; z21 z22];
VC(A*Z)==SP(kron(eye(2),A),VC(Z))                                   % (1)
VC(Z*B)==SP(kron(B',eye(2)),VC(Z))                                  % (2)
VC(C*Z.')==SP(SP(kron(eye(2),C),WIJ(2)),VC(Z))                      % (3)
VC(Z.'*D)==SP(SP(kron(D',eye(2)),WIJ(2)),VC(Z))                     % (4)
VC(A*Z*B+C*Z.'*D)==SP(kron(B',A)+SP(kron(D',C),WIJ(2)),VC(Z)) % (5)
```

4. 计算例7.2.14, 检验两个3 维李代数:

```
% Example 7.2.14 - (1)
M1=[0 0 0 0 0 1 0 -1 0;
    0 0 -1 0 0 0 1 0 0;
    0 1 0 -1 0 0 0 0 0];
SP(M1,WIJ(3)+eye(9))
SPN(M1,M1,eye(27)+WIJ(3,9)+WIJ(9,3))

% Example 7.2.14 - (2)
M2=[0 -7 1 7 0 -2 -1 2 0;
    0 10 -1 -10 0 3 1 -3 0;
    0 -11 2 11 0 -3 -2 3 0];
SP(M2,WIJ(3)+eye(9))
SPN(M2,M2,eye(27)+WIJ(3,9)+WIJ(9,3))
```

索引

《现代数学基础丛书》已出版书目

（按出版时间排序）

1　数理逻辑基础(上册)　1981.1　胡世华、陆钟万　著
2　紧黎曼曲面引论　1981.3　伍鸿熙、吕以辇、陈志华　著
3　组合论(上册)　1981.10　柯召、魏万迪　著
4　数理统计引论　1981.11　陈希孺　著
5　多元统计分析引论　1982.6　张尧庭、方开泰　著
6　概率论基础　1982.8　严士健、王隽骧、刘秀芳　著
7　数理逻辑基础(下册)　1982.8　胡世华、陆钟万　著
8　有限群构造(上册)　1982.11　张远达　著
9　有限群构造(下册)　1982.12　张远达　著
10　环与代数　1983.3　刘绍学　著
11　测度论基础　1983.9　朱成熹　著
12　分析概率论　1984.4　胡迪鹤　著
13　巴拿赫空间引论　1984.8　定光桂　著
14　微分方程定性理论　1985.5　张芷芬、丁同仁、黄文灶、董镇喜　著
15　傅里叶积分算子理论及其应用　1985.9　仇庆久等　编
16　辛几何引论　1986.3　J.柯歇尔、邹异明　著
17　概率论基础和随机过程　1986.6　王寿仁　著
18　算子代数　1986.6　李炳仁　著
19　线性偏微分算子引论(上册)　1986.8　齐民友　著
20　实用微分几何引论　1986.11　苏步青等　著
21　微分动力系统原理　1987.2　张筑生　著
22　线性代数群表示导论(上册)　1987.2　曹锡华等　著
23　模型论基础　1987.8　王世强　著
24　递归论　1987.11　莫绍揆　著
25　有限群导引(上册)　1987.12　徐明曜　著
26　组合论(下册)　1987.12　柯召、魏万迪　著
27　拟共形映射及其在黎曼曲面论中的应用　1988.1　李忠　著
28　代数体函数与常微分方程　1988.2　何育赞　著
29　同调代数　1988.2　周伯壎　著
30　近代调和分析方法及其应用　1988.6　韩永生　著
31　带有时滞的动力系统的稳定性　1989.10　秦元勋等　编著

32 代数拓扑与示性类 1989.11 马德森著 吴英青、段海鲍译
33 非线性发展方程 1989.12 李大潜、陈韵梅 著
34 反应扩散方程引论 1990.2 叶其孝等 著
35 仿微分算子引论 1990.2 陈恕行等 编
36 公理集合论导引 1991.1 张锦文 著
37 解析数论基础 1991.2 潘承洞等 著
38 拓扑群引论 1991.3 黎景辉、冯绪宁 著
39 二阶椭圆型方程与椭圆型方程组 1991.4 陈亚浙、吴兰成 著
40 黎曼曲面 1991.4 吕以辇、张学莲 著
41 线性偏微分算子引论(下册) 1992.1 齐民友 著
42 复变函数逼近论 1992.3 沈燮昌 著
43 Banach 代数 1992.11 李炳仁 著
44 随机点过程及其应用 1992.12 邓永录等 著
45 丢番图逼近引论 1993.4 朱尧辰等 著
46 线性微分方程的非线性扰动 1994.2 徐登洲、马如云 著
47 广义哈密顿系统理论及其应用 1994.12 李继彬、赵晓华、刘正荣 著
48 线性整数规划的数学基础 1995.2 马仲蕃 著
49 单复变函数论中的几个论题 1995.8 庄圻泰 著
50 复解析动力系统 1995.10 吕以辇 著
51 组合矩阵论 1996.3 柳柏濂 著
52 Banach 空间中的非线性逼近理论 1997.5 徐士英、李冲、杨文善 著
53 有限典型群子空间轨道生成的格 1997.6 万哲先、霍元极 著
54 实分析导论 1998.2 丁传松等 著
55 对称性分岔理论基础 1998.3 唐云 著
56 Gelfond - Baker 方法在丢番图方程中的应用 1998.10 乐茂华 著
57 半群的 S - 系理论 1999.2 刘仲奎 著
58 有限群导引(下册) 1999.5 徐明曜等 著
59 随机模型的密度演化方法 1999.6 史定华 著
60 非线性偏微分复方程 1999.6 闻国椿 著
61 复合算子理论 1999.8 徐宪民 著
62 离散鞅及其应用 1999.9 史及民 编著
63 调和分析及其在偏微分方程中的应用 1999.10 苗长兴 著
64 惯性流形与近似惯性流形 2000.1 戴正德、郭柏灵 著
65 数学规划导论 2000.6 徐增堃 著
66 拓扑空间中的反例 2000.6 汪林、杨富春 编著
67 拓扑空间论 2000.7 高国士 著
68 非经典数理逻辑与近似推理 2000.9 王国俊 著
69 序半群引论 2001.1 谢祥云 著
70 动力系统的定性与分支理论 2001.2 罗定军、张祥、董梅芳 编著
71 随机分析学基础(第二版) 2001.3 黄志远 著

72 非线性动力系统分析引论 2001.9 盛昭瀚、马军海 著
73 高斯过程的样本轨道性质 2001.11 林正炎、陆传荣、张立新 著
74 数组合地图论 2001.11 刘彦佩 著
75 光滑映射的奇点理论 2002.1 李养成 著
76 动力系统的周期解与分支理论 2002.4 韩茂安 著
77 神经动力学模型方法和应用 2002.4 阮炯、顾凡及、蔡志杰 编著
78 同调论——代数拓扑之一 2002.7 沈信耀 著
79 金兹堡-朗道方程 2002.8 郭柏灵等 著
80 排队论基础 2002.10 孙荣恒、李建平 著
81 算子代数上线性映射引论 2002.12 侯晋川、崔建莲 著
82 微分方法中的变分方法 2003.2 陆文端 著
83 周期小波及其应用 2003.3 彭思龙、李登峰、谌秋辉 著
84 集值分析 2003.8 李雷、吴从炘 著
85 数理逻辑引论与归结原理 2003.8 王国俊 著
86 强偏差定理与分析方法 2003.8 刘文 著
87 椭圆与抛物型方程引论 2003.9 伍卓群、尹景学、王春朋 著
88 有限典型群子空间轨道生成的格(第二版) 2003.10 万哲先、霍元极 著
89 调和分析及其在偏微分方程中的应用(第二版) 2004.3 苗长兴 著
90 稳定性和单纯性理论 2004.6 史念东 著
91 发展方程数值计算方法 2004.6 黄明游 编著
92 传染病动力学的数学建模与研究 2004.8 马知恩、周义仓、王稳地、靳祯 著
93 模李超代数 2004.9 张永正、刘文德 著
94 巴拿赫空间中算子广义逆理论及其应用 2005.1 王玉文 著
95 巴拿赫空间结构和算子理想 2005.3 钟怀杰 著
96 脉冲微分系统引论 2005.3 傅希林、闫宝强、刘衍胜 著
97 代数学中的 Frobenius 结构 2005.7 汪明义 著
98 生存数据统计分析 2005.12 王启华 著
99 数理逻辑引论与归结原理(第二版) 2006.3 王国俊 著
100 数据包络分析 2006.3 魏权龄 著
101 代数群引论 2006.9 黎景辉 陈志杰 赵春来 著
102 矩阵结合方案 2006.9 王仰贤 霍元极 麻常利 著
103 椭圆曲线公钥密码导引 2006.10 祝跃飞 张亚娟 著
104 凝固过程动力学与交界面稳定性引论 2006.12 徐鉴君 著
105 椭圆与超椭圆曲线公钥密码的理论与实现 2006.12 王学理 裴定一 著
106 非线性演化方程的稳定性与分歧 2007.4 马天 汪宁宏 著
107 正规族理论及其应用 2007.4 顾永兴 庞学诚 方明亮 著
108 组合网络理论 2007.5 徐俊明 著
109 矩阵的半张量积:理论与应用 2007.5 程代展 齐洪胜 著